Taylor & Francis
Taylor & Francis Group

'Feeling' the flow of time through sensorimotor co-ordination

Elio Tuci*, Vito Trianni and Marco Dorigo
IRIDIA, Avenue Franklin Roosevelt 50, CP 194/6 — 1050 Bruxelles, Belgium
email: {etuci,vtrianni,mdorigo}@ulb.ac.be

Abstract. In this paper, we aim to design decision-making mechanisms for a simulated Khepera robot equipped with simple sensors, which integrates over time its perceptual experience in order to initiate a simple signalling response. Contrary to other previous similar studies, in this work the decision-making is uniquely controlled by the time-dependent structures of the agent controller, which in turn are tightly linked to the mechanisms for sensorimotor coordination. The results of this work show that a single dynamic neural network, shaped by evolution, makes an autonomous agent capable of 'feeling' time through the flow of sensations determined by its actions. Further analysis of the evolved solutions reveals the nature of the selective pressures that facilitate the evolution of fully discriminating and signalling agents. Moreover, we show that, by simply working on the nature of the fitness function, it is possible to bring forth discrimination mechanisms that generalize to conditions never encountered during evolution.

Keywords: evolutionary robotics, dynamic neural networks, time-based decision making mechanisms, robustness.

1. Introduction

Animals that forage in a heterogeneous environment, where resources are distributed in patches, are required to make 'complex decisions' such as in which patch to forage, or at which moment in time it is better to move to another patch. To make such decisions, animals need to acquire relevant information from their environment. Although several different mechanisms have been proposed to account for the observed behaviour of different animals, behavioural ecologists tend to assume that the experience the animals have of the patch during time has an incremental or a decremental effect on the animal tendency to remain in the patch (see Nonacs (2003) and Alphen *et al.* (2003) for more on this issue).

A general problem common to biology and robotics concerns the definition of the mechanisms necessary to decide when it is better to pursue a particular action in a certain location and at which moment in time it is better to leave for pursuing a similar or a different activity in a similar or different location. This problem is not limited to foraging alone, but extends to many activities that a natural or artificial agent is required to carry out. Autonomous agents may be asked to change their behaviour in response to the information gained through repeated interactions with their environment. For example, in a group of robots, although many individual actions might be simpler to carry out than a single co-ordinated activity, they might turn out

*Corresponding author.

to be less efficient (see Trianni *et al.*, 2004). Therefore, autonomous agents require adaptive mechanisms to decide whether it is better to pursue solitary actions or to initiate co-operative strategies.

In this paper, we aim at synthesizing decision-making mechanisms for an autonomous robot equipped with simple sensors (see figure 1), which integrates over time its perceptual experiences in order to initiate alternative actions. In other words, the behaviour of the agent should change as a consequence of its repeated interaction with particular environmental circumstances. We are interested in exploiting a biologically-inspired evolutionary approach, based on the use of dynamical neural networks and genetic algorithms (Beer 1995). Generally speaking, the appeal of an evolutionary approach to robotics is twofold. Firstly, and most basically, it offers the possibility of automating a complex design task (Nolfi and Floreano 2000). Secondly, since artificial evolution needs neither to understand nor to decompose a problem in order to find a solution, it offers the possibility of exploring regions of the solution space that conventional design approaches are often constrained to ignore (Harvey *et al.* 1992). In our work, artificial evolution should tightly couple the agent's decision-making mechanisms to the nature of the environment and to the sensorimotor capabilities of the agent.

The experiment performed here, described in detail in section 2, requires an autonomous agent to possess both navigational skills and decision-making mechanisms. That is, the agent should prove capable of navigating in a boundless arena in order to approach a light bulb positioned at a certain distance from its starting position. Moreover, it should prove capable of discriminating between two types of environment: one in which the light can be actually reached, and another in which the light is surrounded by a 'barrier' that prevents the agent from proceeding further toward its target. Due to the nature of the experimental set-up, the agent can find out in which type of environment it is situated only if it proves capable of: (i) moving in a co-ordinated fashion in order to bring forth the perceptual experience required to discriminate between the two environments; and (ii) integrating over time its perceptual experience in order to initiate a signalling behaviour if situated in an environment in which the light cannot be reached.

The results of our simulations show that a single continuous time recurrent neural network (CTRNN, described in section 4.2 and also in Beer (1995)) shaped by evolution makes an autonomous agent capable of 'feeling' time through the flow of sensations determined by its actions.[1] Low-level 'leaky-integrator' neurons, which constitute the elementary units of the robot's controller, provide the agent with the required time-dependent structures. Further analysis of the evolved solutions reveals the nature of the selective pressures that facilitate the evolution of fully discriminating and signalling agents (see section 5). Moreover, we show that, simply by working on the nature of the fitness function, it is possible to bring forth discrimination mechanisms that are robust enough to deal with environmental circumstances that have never been encountered by the best evolved robots' ancestors. This result supports further the significance of the evolutionary robotics approach as a suitable method to develop adaptive autonomous systems.

The paper is structured as follows. Section 2 gives a detailed description of the discrimination task. Section 3 highlights similarities and differences between our approach and some other works in the evolutionary robotics literature about decision-making problems based on the evolution of 'low-level' time-dependent structures. We claim that, although other studies looked at the evolution of time-dependent structures to control the behaviour of agents required to make decisions based on their experience, our experimental set-up allows us to look at issues that have not yet been explored.
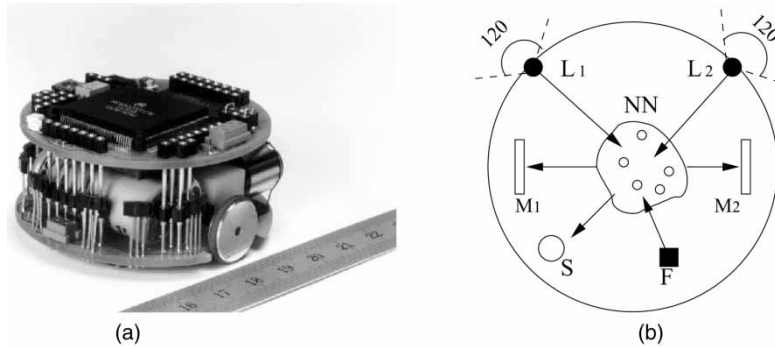
Figure 1. (a) Picture of a Khepera robot. (b) Plan of the robot, showing sensors and motors. The robot is equipped with two ambient light sensors (L_1 and L_2) and a floor sensor indicated by the black square F. The left and right motor (M_1 and M_2) are controlled by a dynamic neural network (NN). A simple sound signalling system, controlled by an output of the network, is referred to as S.

Section 4 introduces the experimental set-up used for the experiments described in sections 5 and 6. Conclusions are drawn in section 7.

2.   Description of the task

At the beginning of each trial, a robot is positioned within a boundless arena, about 100 cm west of a light bulb, with a randomly determined orientation chosen between north-east and south-east (see figure 2, left). The light bulb is always turned on during the trial. The robot perceives the light through its ambient light sensors, positioned 45° left and 45° right with respect to its heading. Light levels alter depending on the robot's distance from the light. The colour of the arena floor is white except for a circular band, centred around the lamp, within which the floor is in shades of grey. The circular band covers an area between 40 and 60 cm from the light; the floor is black at exactly 40 cm from the light; the grey level decreases linearly with the distance from the
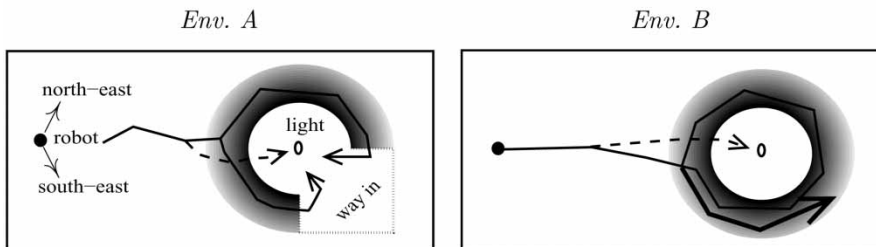

Figure 2. Depiction of the task. The small black circle represents the robot's starting position. The small open circle represents the light bulb. The arena floor is white everywhere except within a circular band surrounding the light. The *way in* zone corresponds to the sector of the band, indicated by dotted lines, in which the floor is white. In both pictures, the continuous arrows are examples of good navigational strategies; the dashed arrows are examples of forbidden trajectories. In Env. B, the continuous arrow gets thicker to indicate that the robot emits a sound after having made a loop around the light.

light. The robot perceives the colour of the floor through its floor sensor, positioned under its chassis, which outputs a value scaled between zero (when the robot is positioned over white floor) and one (when it is over black floor) (see section 4 for a description of the robot).

The robot can move freely within the band, but it is not allowed to cross the black edge. The latter can be imagined as an obstacle or a trough that prevents the robot from further approaching the light (see dashed arrows in figure 2). Whenever the robot crosses the black edge, the trial is unsuccessfully terminated. The area in shades of grey is meant to work as a warning signal, which indicates to the robot how close it is to the danger, i.e. the black edge.

There are two types of environment. In one type—referred to as Env. A—the band presents a discontinuity (see figure 2, left). This discontinuity, referred to as the *way in* zone, is a sector of the band in which the floor is white. In the other type—referred to as Env. B—the band completely surrounds the light (see figure 2, right). The *way in* zone represents the path along which the robot is allowed to reach safely the light in Env. A. A successful robot should prove capable of performing phototaxis as well as looking for the *way in* zone to avoid crossing the black edge of the band. Such a robot should always reach the light in Env. A. On the contrary, in Env. B the robot should, besides avoiding crossing the black edge, signal the absence of the *way in* zone by emitting a tone. How can the robot, provided only with local information, distinguish between environments in which the band presents a discontinuity (i.e. Env. A) and environments in which the band does not presents any discontinuity (i.e. Env. B)?

The cue the agent should use is a temporal one: that is, the Env. B can be 'recognized' by the persistence of a particular perceptual state for the amount of time necessary to discover that there is no *way in* zone. For example, a successful agent might integrate over time the grey level sensed by its floor sensor to bring forth something similar to the 'feeling' of travelling within the band for as long as the time required to complete a loop. Such a strategy would allow the robot to make sure that there is no *way in* zone. Alternatively, the robot might simply react to the colour of the floor and integrate over time the perceived light intensity. In this case, the perception of the circular band is simply used to interrupt the phototaxis and to initiate a circular trajectory.

Notice that, whatever the nature of the perceptual state that the robot integrates over time, the underlying mechanisms for the integration are strongly dependent upon the way in which the robot moves within the environment. For example, let us assume that our robot, by circuiting around the light while remaining over the circular band, integrates over time the reading from the floor sensor. By employing this strategy, the amount of time required for our robot to perform a complete loop of the band depends on the dimensions of the band and on the way in which the robot moves within the band. The robot movements, e.g. its speed and trajectory, are determined by its controller. Thus, the latter should make the robot move in such a way that, if the perception of the band lasts for a certain amount of time, the robot can deduce that the band does not present any discontinuity. Consequently, it should activate the sound signalling.

In view of what we have just said, we claim that the most challenging part of our empirical work resides in: (i) synthesizing, through an evolutionary process, a robot's controller which must be capable of moving the robot co-ordinately so that it can integrate over time the flow of perception determined by its actions; and (ii) evolving within a single, i.e. not modularized, controller the mechanisms required for sensorimotor co-ordination and discrimination through sound signalling.

As illustrated in the next section, the results of previous similar works in the evolutionary robotics literature seem to suggest that CTRNNs provide all the 'building blocks' necessary for evolution to generate the mechanisms required by an autonomous agent to perform this task: that is, mechanisms for sensorimotor co-ordination and time-dependent structures for decision-making (see section 3).

3. Related work

Several studies have described evolutionary simulation models in which time-dependent structures are evolved to control the behaviour of agents required to make decisions based on their experiences. The aim of this section is to highlight similarities and differences between our experiments and those already in the literature.

First of all, we wish to make a distinction between our work and some others that we refer to as non-ecological models (Todd and Miller 1991a,b, Yamauchi and Beer 1994, Tuci *et al.* 2002a). In our model, the agent perception is brought forth by the agent itself through its actions. Contrary to us, in the non-ecological models the perceptual experience of the agents is determined by the experimenter. This is, in our view, a significant difference that bears upon the complexity of our task. Obviously, the flow of perception provides the agents with the cues to make the discrimination. In non-ecological models, the discrimination task is therefore facilitated by the fact that perceptual structures are 'made available' to the agent by the experimenter (see also Parisi *et al.* (1990) for more on this issue). Moreover, some of the non-ecological models (see, e.g. Yamauchi and Beer's (1994) experiment) are further simplified by the presence of an explicit reinforcement signal, i.e. an input signal explicitly dedicated to informing the agent's controller of the characteristics of the 'environmental circumstances' in which it is currently situated by making available to the system any possible mismatch between the current agent's action and the correct response.

Other studies on the evolution of time-dependent structures for discrimination tasks share with our experiment the ecological perspective, in which the nature of the agent's perception is determined by its own actions and the reinforcement signals are part of the evolved structures (see Ziemke and Thieme 2002, Tuci *et al.* 2002b, Nolfi 2002, Blynel and Floreano 2003). The evolution of time-dependent structures and decision-making mechanisms has been studied extensively on the T-maze problem (see Ziemke and Thieme 2002, Blynel and Floreano 2003). Generally speaking, these tasks require a robot to find its way to a goal location placed at the bottom of any of the two arms of a maze. When at the T-junction, the robot must decide whether to turn left or right. The correct decision can be made if the agent is capable of exploiting perceptual cues that were available to it while it was navigating down the first corridor, or by 'remembering' something about previous trials in a similar T-maze. In Ziemke and Thieme (2002), a mechanism for neuromodulation of sensorimotor weights provided the agents with the required plasticity to exploit the relationship between the location of light signals placed roughly at the middle of the first corridor and the turn to make at the junction. Blynel and Floreano (2003) allowed the agent to experience the environment in a first trial, in which success or failure play the role of a reinforcement signal, in order to associate the position of the goal with respect to the T-junction. In Tuci *et al.* (2002b), evolved CTRNNs provided the agents with the required plasticity to discover the spatial relationship between the position of a landmark and the position of a goal. In this study, the spatial relationship between the goal and the landmark can be learned by 'remembering' from previous trials the relative position of the landmark with respect to the goal.

The difference between the ecological models and our study is not as apparent as it was for the non-ecological ones described at the beginning of the section. However, it should be noticed that, in the ecological studies reviewed above, the discrimination is based on the recognition of distinctive environmental contingencies and the maintenance of these experiences through time, as a form of short-term memory.[2] On the contrary, in our study the cue that allows the agent to make the discrimination has to deal with the persistence over time of a perceptual state common to both the elements to be distinguished, i.e. Env. A and Env. B, rather than with the nature of the cue itself employed to make the discrimination. That is, in our case, due to the nature of the agent sensory apparatus, one environment can be distinguished from the other solely because a perceptual state, common to both environments, might, in one case, be perceived by the agent for a time longer than what the agent might experience by acting in the other type of environment.

Experiments similar to the one described here were performed by Nolfi (2002) and by Croon *et al.* (2004). These authors investigated a discrimination task in which a robot, while navigating through a maze, must recognize that it is located in one room rather than in another. In spite of the differences in the experimental set-up, these works and the one described here focus on similar issues. They all exploit evolution to design controllers for autonomous robots required to make decisions based on time-dependent structures.

4. The robot-based model

In this section, we provide some details concerning the robot-environment simulation model used to evolve the controllers (see section 4.1), the equation used to update the state of the neural network (see section 4.2) and the parameters of the genetic algorithm (see section 4.3).

4.1. *The simulation*

The robot and its world are simulated using a modified version of the 'minimal simulation' technique described by Jakobi (1997). Jakobi's technique uses high levels of noise to guarantee that the simulated controller will transfer to a physically realized robot with no loss of performance. Our simulation models a Khepera robot, a 55-mm diameter cylindrical robot (see figure 1). This simulated robot is provided with two ambient light sensors, placed at $45°$ (L_1) and $-45°$ (L_2) with respect to its heading, and a floor sensor positioned facing downward on the underside of the robot (F). The light sensors have an angle of acceptance of $120°$, and they can detect the light up to a distance of 100 cm from the light source. Light levels change as a function of the robot's distance from the lamp. The light sensor values are extrapolated from a look-up table[3] which corresponds to the one provided with the Evorobot simulator—see Nolfi (2000) for further details. The floor sensor can be conceived of as a proximity infrared sensor capable of detecting the level of grey of the floor. It produces an output that is proportional to the level of grey, scaled between zero (when the robot is positioned over white floor) and one (when it is over black floor). The sound signalling system is represented by the binary output of one of the neurons of the robot's controller (see section 4.2 for details).

The robot has right and left motors—respectively M_1 and M_2—which can move independently forward or backward, allowing it to turn fully in any direction. The robot is assumed to have negligible mass, so that the motor output can be taken as the tangential velocity of the robot at the motor mount point. The current heading θ' and the position

(x', y') of the virtual Khepera within its environment are calculated at each time step according to the following equations:

$$\mu = r_w * \omega_w * \Delta t;$$

$$x' = x + \chi(\theta) * \frac{(\omega_r + \omega_1)}{2} * \mu * \mathrm{d}t;$$

$$y' = y + \gamma(\theta) * \frac{(\omega_r + \omega_1)}{2} * \mu * \mathrm{d}t;$$

$$\theta' = \theta + \frac{(\omega_r - \omega_1) * \mu}{d * \mathrm{d}t}, \tag{1}$$

where ω_r and ω_1 correspond to the instantaneous velocity of the right and the left motors, respectively; μ is a constant equal to 0.8 cm representing the maximum distance covered by a robot in $\Delta t = 1$ s at maximum speed $\omega_w = 10$ rad s^{-1}; $r_w = 0.8$ cm is the the radius of the wheel; $\chi(\theta)$ and $\gamma(\theta)$ correspond to the horizontal and vertical increments for a Khepera travelling at a speed of 1 cm per second for a given orientation θ; and $d = 5.2$ cm is the length of the axle. The simulation time step $\mathrm{d}t$ corresponds to 0.2 s. There is no allowance for momentum, and the noise inherent in the real-world situation is not modelled.

4.2. *The controller*
Fully connected, eight neuron CTRNNs are used. All neurons are governed by the following state equation:

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \frac{1}{\tau_i}\left(-p_i + \sum_{j=1}^{8} w_{ji}\sigma(p_j + \beta_j) + gI_i\right), \quad \sigma(x) = \frac{1}{1 + e^{-x}}, \tag{2}$$

where, using terms derived from an analogy with real neurons, p_i represents the cell potential, τ_i the decay constant, β_j the bias term, $\sigma(p_j + \beta_j)$ the firing rate, w_{ji} the strength of the synaptic connection from neuron jth to neuron ith, and I_i the intensity of the sensory perturbation on sensory neuron i. Three neurons receive input I_i from the robot sensors. These input neurons receive a real value in the range [0,1], which is a simple linear scaling of the reading taken from its associated sensor.[4] The other neurons do not receive any input from the robot's sensors. The cell potential p_i of the sixth neuron, mapped into [0,1] by the sigmoid function σ and then set to unity if bigger than 0.5 or zero otherwise, is used by the robot to control the sound signalling system. The cell potentials p_i of the seventh and the eigth neurons, mapped into [0,1] by the sigmoid function σ and then linearly scaled into $[-10,10]$ set the robot motors output. The strength of synaptic connections w_{ji}, the decay constants τ_i, the bias terms β_j and the gain factor g are genetically encoded parameters. Cell potentials are set to zero any time the network is initialized or reset, and circuits are integrated using the forward Euler method with an integration step size of 0.2 s.

4.3. *The evolutionary algorithm*
A simple generational genetic algorithm (GA) is employed to set the parameters of the networks (Goldberg 1989). The population contains 100 genotypes. Generations following the first one are produced by a combination of selection with élitism, recombination and mutation. For each new generation, the three highest scoring individuals ('the élite')

from the previous generation are retained unchanged. The remainder of the new population is generated by fitness-proportional selection from the 70 best individuals of the old population. Each genotype is a vector comprising 81 real values (64 connections, eight decay constants, eight bias terms and a gain factor). Initially, a random population of vectors is generated by initializing each component of each genotype to values chosen uniformly random from the range [0,1]. New genotypes, except 'the élite', are produced by applying recombination with a probability of 0.3 and mutation. Mutation entails that a random Gaussian offset is applied to each real-valued vector component encoded in the genotype, with a probability of 0.15. The mean of the Gaussian is zero, and its standard deviation is 0.1. During evolution, all vector component values are constrained to remain within the range [0,1]. Genotype parameters are linearly mapped to produce CTRNN parameters with the following ranges: biases $\beta_j \in [-2,2]$, weights $w_{ji} \in [-6,6]$ and gain factor $g \in [1,12]$. The genes that codify the decay constants are firstly linearly mapped on to the range $[-0.7, 1.7]$ and then exponentially mapped into $\tau_i \in [10^{-0.7}, 10^{1.7}]$.

5. Evolution of time-dependent discrimination mechanisms

In this section, we illustrate the fitness function and the results of a first series of experiments in which we aim to evolve agents capable of discriminating between Env. A and Env. B. The fitness function employed simply rewards a robot for approaching the light bulb, and for signalling anytime it is located in Env. B, but not signalling in Env. A. A significant feature of this fitness function is that it rewards agents that make a proper use of their sound signalling system regardless of the strategies exploited to make the discrimination.

We have run three sets of 10 evolutionary simulations—referred to as condition +A, condition +B and condition AB—which differ with respect to the proportion of Env. A and Env. B each agent experiences during the evolutionary phase. In condition +A, the Env. A are three times more frequent than the Env. B. In condition +B, the Env. B are three times more frequent than the Env. A. In condition AB, the two types of environment appear with the same frequency.

The reason for running simulations on these three conditions is related to the potential effects on the phylogeny of the system produced by different selective pressures determined by the proportion of Env. A and Env. B the agents encounter during their lifetime. For example, due to the nature of the task, we may assume that sound signalling is advantageous only if it is employed by an agent that possesses the sensorimotor co-ordination required to discriminate between Env. A and Env. B. However, if Env. B is more frequent than Env. A—as in condition +B—an agent that always signals, regardless of its sensorimotor capabilities, has a higher fitness than an agent that never signals. Although both types of agent are sometimes right and some other times wrong, the former is three times more successful than the latter. Thus, in condition +B, evolution might progress through an initial stage characterized by populations of signalling agents, followed by a subsequent stage characterized by populations of agents that combine signalling with the sensorimotor co-ordination required to carry out the discrimination task.

On the contrary, in condition +A the pay-off is reversed. An agent that never signals is three times more successful than an agent that always emits a sound signal. Thus, evolution might proceed by firstly rewarding agents capable of sensorimotor co-ordination but not capable of sound signalling, and subsequently by rewarding those agents that combine sensorimotor co-ordination with a proper use of the sound. Both conditions +A and +B require the gradual and progressive evolution of adaptive mechanisms

coupled with previously evolved traits. However, such evolutionary trends might not be particularly easy to obtain.

It is more difficult to imagine what could happen in condition AB, which neither favours the mechanisms for sensorimotor co-ordination over the mechanisms for signalling, nor vice versa. Although this condition might facilitate the progressive evolution of agents capable of the required sensorimotor co-ordination and signalling, there is also the possibility that the lack of a strong selective pressure for or against either mechanism might be deleterious rather than beneficial to the evolution of the desired behaviour. The system might end up fluctuating without being capable of 'bootstrapping' from randomly initialized controllers to the desired solution/s.

The results of the simulations and the analysis of the evolved solutions, illustrated in section 5.2, provide useful empirical evidence to clarify some of the issues discussed above.

5.1. *The evaluation function*

During the evolution, each genotype is coded into a robot controller and is evaluated 16 times. The proportion of Env. A and Env. B within the 16 trials depends on the condition. At the beginning of each trial, the neural network is reset, i.e. the activation value of each neuron is set to zero. Each trial differs from the others in the initialization of the random number generator, which influences the robot starting position and orientation, the position of the *way in* zone, and the noise added to motors and sensors. For each trial in Env. A, the position of the *way in* zone is varied to facilitate the evolution of robust navigational strategies. Its amplitude is fixed to $\pi/2$. Within a trial, the robot life-span is 80 s (400 simulation cycles). A trial is terminated earlier if either the robot crosses the black edge of the band (see dashed arrows in figure 2) or because it reaches a Euclidean distance from the light higher than 120 cm. In each trial e, the robot is rewarded by an evaluation function f_e which corresponds to the sum of the following two components:

(1) R_{motion}: this component rewards movements toward the light bulb, and it is computed as:

$$R_{\mathrm{motion}} = \frac{d_{\mathrm{f}} - d_{\mathrm{n}}}{d_{\mathrm{f}}}, \tag{3}$$

where d_{f} and d_{n} represent, respectively, the furthest and the nearest Euclidean distance between the robot and the light bulb. In particular, d_{f} is updated whenever the robot increases its maximum distance from the light bulb. At the beginning of the trial, d_{n} is fixed as equal to d_{f}, and it is subsequently updated: (i) every time the robot gets closer to the light bulb; (ii) every time df is updated. In this latter case, d_{n} is set up equal to the new d_{f}. In Env. A, d_{n} is set to zero if the robot is less than 7.5 cm away from the light bulb. In Env. B, d_{n} is set to zero as soon as the robot reaches the band in shades of grey.

(2) R_{signal}: this component rewards agents that: (i) do not signal any time they are located in Env. A; and (ii) emit a sound signal any time they are located in Env. B. The component is computed as:

$$R_{\mathrm{signal}} = \begin{cases} 1 & \text{if proper signalling} \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

An important feature of this evaluation function is that it simply rewards agents that make a proper use of their sound signalling system, without directly interfering with the nature of the discrimination strategies.

5.2. *Results obtained*

Ten replications of the experiments are run for each condition. Figure 3 shows, for each condition, the fitness of the best individual and the mean population fitness plotted against the generation number and averaged over the 10 replications. It is possible to notice that in all the three conditions, the maximum fitness value is reached suggesting that in all replications of the experiment and in all conditions a successful behaviour was evolved. The 100% success rate can be accounted for by recalling that the fitness function, not rewarding any specific action except phototaxis and the signalling behaviour, has positively influenced the development of successful behaviours. In fact, evolution is left free to search for whatever strategy could be effective for the achievement of the final goal.[5]
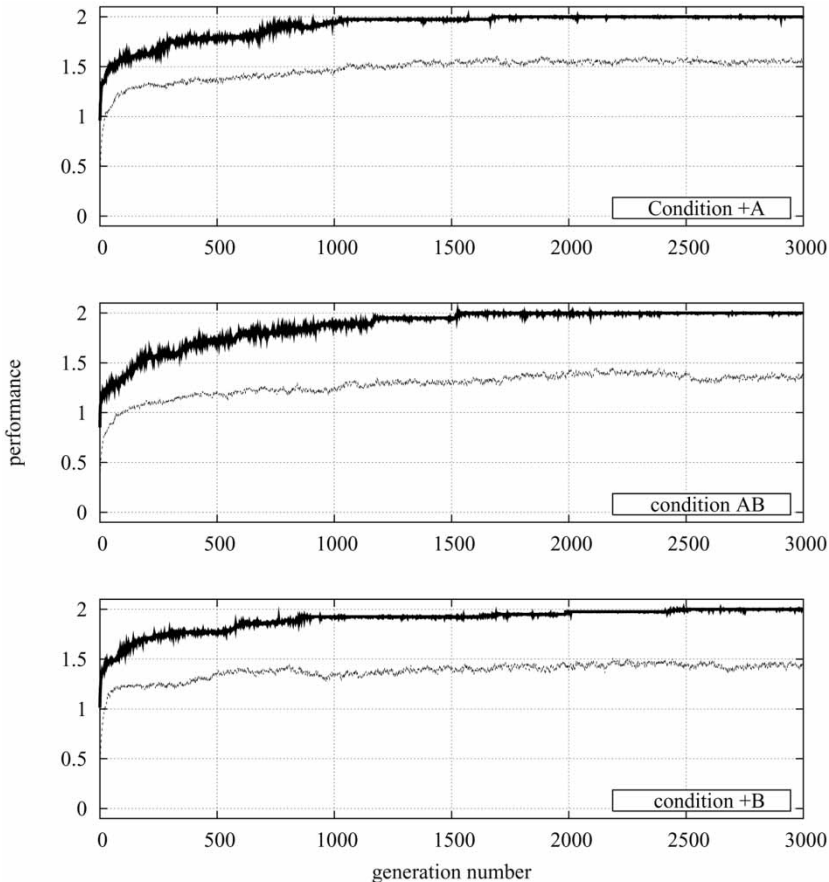

Figure 3. Average fitness during the evolution. All plots are the average over the 10 replications of the experiment performed for each condition. The thick line correspond to the average fitness of the best individual, while the thin line refers to the average fitness of the population.

5.2.1. *Analysis of the evolved behavioural strategies.* A qualitative analysis of the evolved controllers confirms that a number of different behavioural strategies have been obtained. However, some constant characteristics can be recognized. At the beginning of a trial, all robots perform phototaxis until they reach the circular band. When the grey level on the floor exceeds a certain threshold, the robots start circuiting around the light bulb with an approximately constant angular speed. Whenever the robots are placed in Env. A and the *way in* zone is detected, phototaxis starts again and the light bulb is reached. On the contrary, in Env. B, after travelling on the band for a given time without detecting the *way in* zone, the robots initiate a signalling behaviour.

An example of this behaviour is shown in figure 4: in both Env. A and Env. B, it is possible to see that, when the circular band is detected—see continuous line F at about simulation cycle 130—the robot starts moving on the circular band, maintaining a constant distance from the light bulb. This behaviour is indicated by the constant readings of the light sensors L1 and L2 and of the floor sensors F. In Env. A, the *way in* zone is encountered shortly before simulation cycle 300, as indicated by the sudden drop in the floor sensor F. At this point, the robot performs phototaxis again, rapidly reaching the light bulb, as indicated by the high activation of the light sensors L1 and L2 at the end of the simulation.

The constant angular speed on the circular band is the basic mechanism exploited for discrimination between Env. A and Env. B by successfully evolved robots. In fact, this constant motion allows the robots to experience a constant perceptual state (the grey level of the floor and the light intensity that impinges on their sensors), which roughly corresponds to the constant flow of time. In figure 4, it is possible to see that the persistence of a particular perceptual state, corresponding to the robot circuiting around the light and over the band, makes the output S, which controls the sound, increase linearly. This perceptual state triggers the sound signalling through an efficient integration mechanism which is based on the 'feeling' of travelling long enough over the circular band without having encountered the *way in* zone. In fact, if the *way in* zone is encountered, as in the upper part of figure 4, the activation of the neuron S decreases below the threshold level 0.5. This response makes the robot capable of avoiding initiating the signalling behaviour when it is not required. The situation is different in Env. B: the absence of the *way in* zone let the output of neuron S reach and overcome the threshold level 0.5—see bottom part of figure 4, simulation cycle 300. This response makes the robot capable of correctly signalling that it is located in an Env. B.

In summary, the behavioural analysis revealed that the evolved controllers produce the required sensorimotor co-ordination that brings forth a perceptual state that is integrated over time and exploited for discrimination through sound signalling.

5.2.2. *Analysis of the different selective pressures.* All the replications produced a successful controller, no matter the condition in which the evolution took place (+A, +B or AB). Therefore, we still cannot state if the selective pressures introduced in condition +A, in condition +B and in condition AB made any significant difference. From figure 3, it is possible to see that in condition +A the average performance of the best individual reaches the optimum earlier than in the other two cases. This seems to suggest that the evolutionary regime in condition +A might be more suitable for the evolution of fully discriminating agents than the one corresponding to condition +B and condition AB. Recall that the condition +A facilitates the early evolution of a population of agent with the required sensorimotor co-ordination followed by population of agents capable of successfully discriminating through sound signalling.

Environment type A
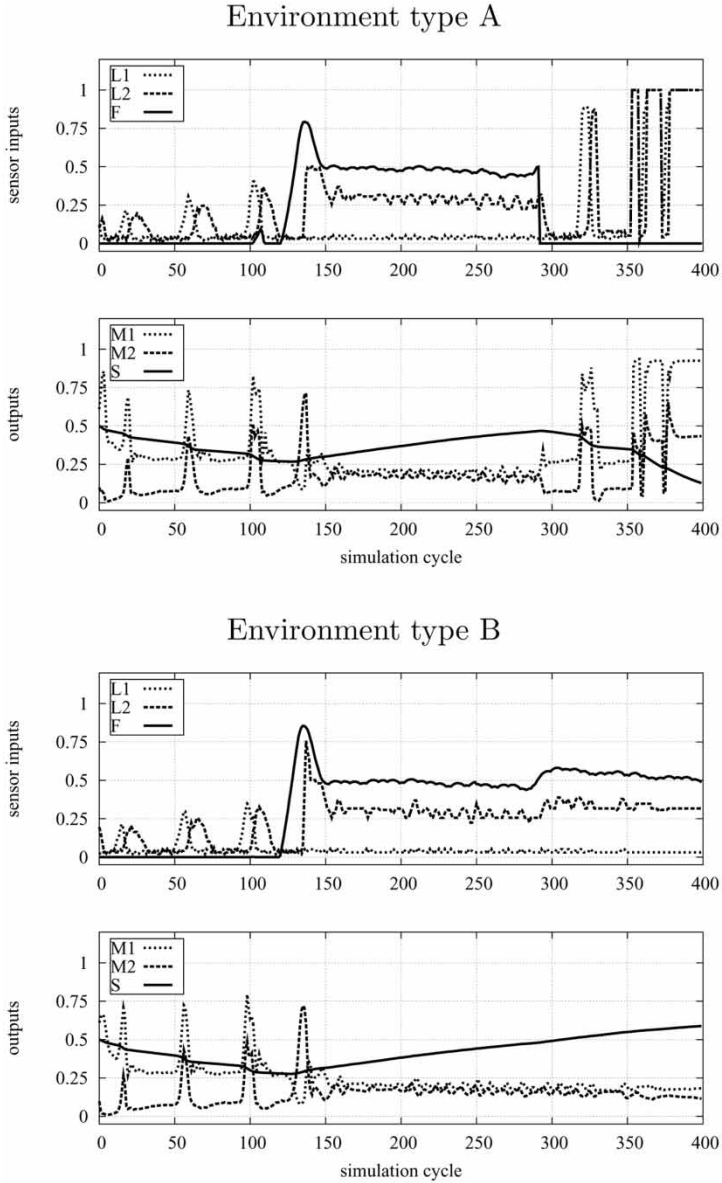

Environment type B

Figure 4. Behavioural analysis. The sensors' activity and the corresponding motor outputs are plotted for 400 simulation cycles. L1 and L2 refer to the light sensors, while F refers to the floor sensor. M1 and M2 correspond to the motors of the two wheels, and S refers to the sound signalling. When S is bigger than 0.5, the robot emits a signal (see section 4.1).

In order to assess and compare the performance of controllers evolved in different conditions, we performed further analyses by re-evaluating each of the best evolved final generation individuals for 100 trials in each type of environment (i.e. Env. A and Env. B). In each trial performed in Env. A, we look at the robot capability of reaching the light bulb (*Succ.*), without making any error. Errors can be of two types: $E1$ refers to the emission of a sound signal, while $E2$ refers to crossing the black edge of the band. Similarly, in Env. B, we look at the performance of the robot on properly signalling

the absence of the *way in* zone (*Succ.*), without committing any error. Also in this case, two error types are possible: $E3$ refers to the lack of sound signalling, and $E4$ refers to the robot crossing the black edge of the band. Furthermore, in Env. B we also compute the offset between the entrance position of the robot in the circular band and the position in which the robot starts to signal. This measure, called offset Δ, is computed as follows:

$$\Delta = |\alpha(t_e, t_s)| - 2\pi, \tag{5}$$

$$\alpha(t_1, t_2) = \sum_{t=t_1}^{t_2-1} \widehat{\mathbf{AOB}}, \quad \mathbf{A} = \mathbf{X}_t, \mathbf{B} = \mathbf{X}_{t+1}, \tag{6}$$

where $\mathbf{O}$ corresponds to the position of the light, and α is the angular displacement of the robot around the light from the starting position—the position at time t_e when the robot enters into the circular band—to the signalling position—the position at time t_s when the robot starts signalling. α is computed summing up all the convex angles $\widehat{\mathbf{AOB}}$ comprised between two consecutive position of the robot $\mathbf{X}_t$, taking into account that an angle is negative if the robot moves clockwise. This measure accounts for the capability of a robot for searching the way in zone. Offset Δ takes value zero if the robot signals exactly after covering a complete loop of the circular band. Otherwise, it gives the angular displacement from this position. Negative values of the offset Δ suggest that the robot signals before having performed a complete loop, while positive values correspond to the situation in which the robot has performed more than one loop around the light, waiting too long to signal.

Table 1 refers to the post-evaluation in condition $+$A. Here, all the evolved controllers perform well, having a very high success rate in both Env. A and Env. B. It is worth

Table 1. Post-evaluation in condition $+$A. Performance of the best evolved controllers of each replication. The percentage of success (*Succ. %*) and the percentage of errors ($E1$ and $E2$ in Env. A, and $E3$ and $E4$ in Env. B) over 100 trials are shown for both Env. A and Env. B. Additionally, the average offset Δ and its standard deviation (degrees) are shown for Env. B.

	Condition $+$A							
	Env. A			Env. B				
							Offset Δ	
Run number	*Succ.* (%)	*E1* (%)	*E2* (%)	*Succ.* (%)	*E3* (%)	*E4* (%)	Average	Standard
1	100	0	0	100	0	0	−38.5	8.79
2	100	0	0	99	1	0	−60.05	30.47
3	100	0	0	100	0	0	−57.47	12.6
4	100	0	0	99	0	1	−17.94	24.06
5	91	1	8	90	0	10	−67.21	25.78
6	100	0	0	98	2	0	−28.83	38.38
7	98	0	0	100	0	0	−47.16	25.21
8	97	0	3	100	0	0	−65.49	16.04
9	96	0	4	91	0	9	63.98	22.91
10	98	0	2	96	4	0	−57.47	27.5

noting that there are only few cases in which the robot makes signalling errors ($E1$ and $E3$), while some replications of the experiments have a higher error rate in crossing the black edge of the circular band. This is due mainly to a tendency of the robots to approach the black edge while circuiting on the band. Concerning the offset Δ, most evolved controllers have a negative value, in general lower than $65°$, meaning that all robots signal far before having completed one loop of the circular band. However, this offset is enough to discriminate between Env. A and Env. B, as the way in zone is $90°$ wide. Only in one case, in replication 9, is the robot 'prudent': that is, it signals only after having completed a loop around the light bulb.

Here, the selective pressure given by the higher percentage of Env. A encountered by the robot during evolution yields a robust behaviour. In condition +A, the sound signalling behaviour appears only after having acquired the sensorimotor co-ordination required for the integration over time.[6] Therefore, it is simpler to make the association between the sound signalling behaviour and the absence of the *way in* zone with respect to the other conditions, as we shall see in the following.

The situation in condition +B is completely reversed, as shown in table 2. As expected, the robots perform well in Env. B, with a high success rate and low percentage of both types of error. However, $E1$, that is, signalling when not required, is much higher than in condition +A.[7] In all these cases, robots trigger the sound signal in Env. A too early, before finding the *way in* zone. This is confirmed by the wide offset Δ recorded in Env. B. The selective pressure given by the higher percentage of Env. B first favoured the evolution of signalling robots, no matter the environment in which they were located and their sensorimotor capabilities. The results obtained suggest that, in condition +B, it is more difficult to evolve agents capable of moving in a co-ordinated fashion in order to discriminate successfully, through sound signalling, Env. A from Env. B.

The results of the post-evaluation in condition AB, shown in table 3, are intermediate with respect to the other two conditions. While a high success rate is achieved for many of the best evolved robots, we can also observe some errors, performed mainly in

Table 2. Post-evaluation in condition +B; see caption of table 1 for details.

	Condition +B							
	Env. A			Env. B				
							Offset Δ	
Run number	*Succ.* (%)	$E1$ (%)	$E2$ (%)	*Succ.* (%)	$E3$ (%)	$E4$ (%)	Average	Standard
1	97	0	3	100	0	0	−58.49	6.28
2	86	13	1	100	0	0	−166	13.58
3	100	0	0	99	1	0	−63.63	33.69
4	92	8	0	100	0	0	−59.30	18.16
5	84	13	0	100	0	0	−161.65	13
6	78	12	10	100	0	0	−125.08	7.44
7	75	24	1	100	0	0	−181.23	31.51
8	62	36	1	98	2	0	−196	83.07
9	68	30	2	100	0	0	−173.78	21.77
10	98	0	0	100	0	0	65.70	50.42

Table 3. Post-evaluation in condition AB; see caption of table 1 for details.

	Condition AB							
	Env. A			Env. B				
							Offset Δ	
Run number	*Succ.* (%)	E1 (%)	E2 (%)	*Succ.* (%)	E3 (%)	E4 (%)	Average	Standard
1	90	9	1	100	0	0	−112.48	25.21
2	98	0	2	100	0	0	−1.32	26.35
3	86	0	14	100	0	0	−54.61	5.15
4	100	0	0	100	0	0	28.46	8.02
5	96	2	2	99	0	1	−66.64	33.23
6	99	0	1	100	0	0	−29.40	36.66
7	89	0	11	99	0	1	85.76	45.26
8	90	0	5	96	4	0	55.96	13.17
9	97	0	3	100	0	0	−56.9	30.36
10	96	0	4	96	0	4	70.86	82.50

Env. A. The robot run number 1 signals too early—as indicated by the high negative offset. Thus, it makes many errors of type $E1$. A similar performance was produced by several other controllers evolved in condition +B. Other robots perform well in sound signalling, but they often cross the black edge of the circular band, making errors $E2$ or $E4$. This time, a similar performance was produced by several others controllers evolved in condition +A. Clearly, the condition AB has a more uncertain and variable outcome than the other two.

In conclusion, the above analysis reveals that the selective pressure corresponding to condition +A, which favours the progressive evolution initially of sensorimotor co-ordination followed by discrimination capabilities through sound signalling, produces an overall better performance of the evolved controllers.

6. Evolution of robust discrimination strategies

One of the most desirable features of an autonomous robot consists in its capability to adapt to varying environmental conditions. Research in biology has shown that the 'adaptability' of natural systems is generally bound within some limits that are determined by the evolutionary history of the species (see Breland and Breland 1961, Bolles 1970, Johnston 1981). Animals, especially humans, are also capable of coping with circumstances that it is reasonable to assume they have never encountered during their phylogeny. Although the nature of the mechanisms that allow an animal to adapt to 'novel' circumstances is still a controversial issue, biologists tend to assume that mechanisms evolved to accomplish a particular task might subsequently carry out other adaptive functions different from those for which they have been evolved (Gould and Vrba 1982).

Evolutionary robotics makes it possible to bring these concepts within the domain of artificial autonomous systems (see Wheeler 1996). That is, by evolving robot's controllers in varying environmental conditions, we might obtain agents that show the following characteristics: (i) they can easily adapt to several environmental conditions as long as they have experienced them during evolution; and (ii) a phylogenetic history in varying

environments might help the emergence of mechanisms that allow the final generation agents to cope with circumstances that their ancestors have never encountered.

In this section, we show some results in which the robot's controllers are evolved in environments in which the distance between the black edge of the band and the light bulb varies from a minimum of 20 cm to a maximum of 65 cm. We refer to this distance as light–band distance, and the set of simulations in which this distance varies as condition *R*. The environmental variation experienced by the robots in condition *R* represents a significant evolutionary challenge for robots that are required to perform the discrimination task described in section 2. By varying the light–band distance, while maintaining fixed the width of the circular band (20 cm), the spatiotemporal structures that the robot must exploit to distinguish between Env. A and Env. B vary as well. For example, for a robot that moves at a certain speed and with a certain trajectory over the band, if the light–band distance is at its minimum of 20 cm, the time required to perform a loop around the light will be definitely shorter than the time required in an environment in which the light–band distance is at its maximum of 65 cm. In order to be capable of successfully distinguishing between Env. A and Env. B, this robot must be able to adapt to the characteristics of the environment. One possible solution consists of exploiting the relationship between the intensity of the light and the activation of the floor sensor in order to 'get an idea' of the light–band distance. If the robots can perceive how far the light is from the black edge of the band, they can use this to estimate the inner circumference of the circular band. They can then 'deduce' the length of time for which they need to circuit the circular band before signalling. In principle, once evolved, this adaptive mechanisms should be robust enough for the robot to cope with environmental circumstances never encountered by its ancestors.

Unfortunately, preliminary tests in which populations of robots have been evolved in varying environmental circumstances with the fitness function described in section 5 were not particularly satisfying. The variation of important characteristics of the environment was not enough to evolve adaptive agents capable of adjusting, through the exploitation of the light–band distance, their discrimination strategies to the current environmental circumstances. An adaptive discrimination strategy would require the robot to signal after having made a loop around the light while remaining on the circular band. Evolution found a simpler solution: whatever the light–band distance was, the robots were simply circuiting around the light for as long as the time required to trigger the sound signalling. In more detail, the evolved robots started signalling after a fixed amount of time, chosen in a way to guarantee that they would start signalling after having made a loop around the light in those environments in which the light–band distance was at its maximum. Obviously, this strategy allows them to make perfect discrimination also in those environments in which the light–band distance was shorter. However, the shorter the light–band distance, the more loops around the light were required before emitting the sound signal, and the less efficient the robot's discrimination strategy appears. Moreover, by employing this strategy, the robots were sometimes failing to discriminate properly if located in environments in which the light–band distance was higher than the maximum distance experienced by their ancestors during evolution. That is, by employing a fixed 'signalling time', the robots were erroneously signalling in environments in which there was a *way in* zone.

Thus, we have designed a new fitness function that rewards discriminating robots for signalling just after having made a loop around the light bulb (see section 6.1 for a detailed description of the fitness function). Recall that the fitness function employed in previous simulations was simply rewarding agents for properly signalling regardless

of the strategies employed to make the discrimination. By rewarding robots that signal after having only made a single loop around the light in an Env. B, we are expecting that successful agents should evolve adaptive mechanisms that allow them to find a way to perceive the light–band distance. Hopefully, such robots should be able to adjust the time required for signalling with respect to this distance.

6.1. The evaluation function

During the evolution, each genotype is coded into a robot controller, and is evaluated 16 times in five different environments, i.e. 80 trials in total. The five environments differ as far as concerns the light–band distance. For each environment, this distance varies randomly within one of the following intervals: [20, 25] cm, [30, 35] cm, [40, 45] cm, [50, 55] cm, [60, 65] cm. Within a set of 16 trials, the robots experience the same proportion of Env. A and Env. B.

In each trial, the initializations of the robot's controller and of the random parameters of the evaluation are performed in the same way as explained in section 5.1. Within a trial, the robot life-span is 120 s (600 simulation cycles). As explained before, a trial is terminated earlier if either the robot crosses the black edge of the band (see dashed arrows in figure 2) or because it reaches a Euclidean distance from the light higher than 120 cm. In each trial e, the robot is rewarded by an evaluation function f_e, which corresponds to the sum of two components: R_{motion}, which rewards phototaxis and is computed as in equation (3); and R_{signal}, which accounts for proper signalling of the robot in Env. A and in Env. B. R_{signal} is computed as illustrated in equation (4) when the robot is placed in Env. A. Otherwise, we take into account the signalling error the robot makes with respect to the completion of a loop around the light:

$$R_{\text{signal}} = \begin{cases} 1 - \dfrac{|\Delta|}{2\pi} & \text{proper signalling} \\ 0 & \text{otherwise} \end{cases}, \tag{7}$$

where the offset Δ was introduced in equation (5). If we assume that a whole loop around the light bulb while remaining over the circular band is needed for the robot to make sure that there is no *way in* zone, then R_{signal} rewards adaptive signalling strategies, which adjust themselves to the current light–band distance.[8]

6.2. *Results obtained*

The condition R is studied performing 10 evolutionary runs, using the experimental set-up described earlier. Figure 5 reports the fitness of the best individual and the mean fitness of the population during the evolution, averaged over the 10 replications. The maximum fitness is reached in all replications, as indicated by the plot in figure 5. However, the performance of the best individual oscillates near the maximum value, due to the high variability in the fitness component R_{signal}. As expected, it is quite difficult for the robots to signal exactly after having covered one loop around the light.

A post-evaluation analysis has been performed in order to assess the discrimination capabilities of the final generation best evolved controllers in each of the 10 replications of the experiment. We have re-evaluated each of these controllers 200 times in each of the five environmental conditions—100 trials in Env. A and 100 trials in Env. B. The results, averaged for all the replications and for the five environmental circumstances, are summarized in table 4. This table illustrates, for each agent, two measures of its
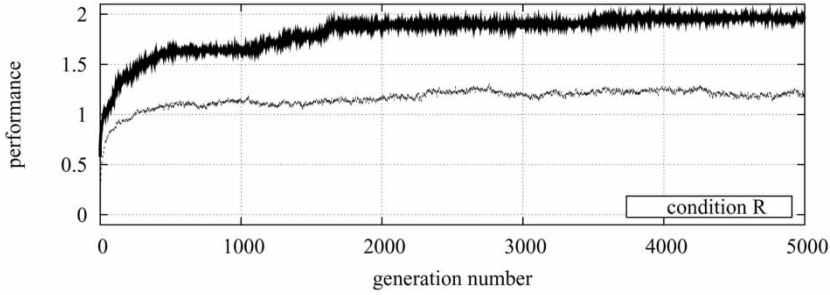
Figure 5. Average fitness during the evolution. The plot corresponds to the average over the 10 replications of the experiment. The thick line corresponds to the average fitness of the best individual, while the thin line refers to the average fitness of the population.

average success rate (see columns R_{motion} and R_{signal}), and two measures of its average rate of failure for each type of environment (see columns $E1$, $E2$, $E3$ and $E4$). Table 4 shows that, as far as concerns both R_{motion} and R_{signal}, each final generation best evolved controller possesses a very high success rate for both Env. A and Env. B. In particular, the performance in Env. A is always very close to the maximum, and very few errors have been observed. As far as concerns Env. B, it is important to highlight the good performance in the R_{signal} component, which is related to the offset Δ in the signalling behaviour—see equation (5). Recall that a value of 0.9 for R_{signal} corresponds to an absolute mismatch of $36°$ with respect to a complete loop. This means that the evolved signalling behaviour is very accurate, despite the high variability (i.e. the light–band distance) in the environmental conditions encountered by the robot.

The evolved controllers produce behaviours that are qualitatively similar to those described in the previous section. As it was for the robots evolved in the other evolutionary conditions (i.e. +A, +B, AB), also for the best evolved robot in condition

Table 4. Condition R. Performance of the best evolved controllers of each replication. R_{motion}, R_{signal} were introduced in sections 5.1 and 6.1. $E1$, $E2$, $E3$ and $E4$ were introduced at the beginning of section 5.2.2.

	Condition R							
	Env. A				Env. B			
Run number	R_{motion}	R_{signal}	$E1$ (%)	$E2$ (%)	R_{motion}	R_{signal}	$E3$ (%)	$E4$ (%)
1	0.99	0.98	0	1	1	0.91	5	23
2	1	1	0	0	1	0.97	0	0
3	0.99	0.99	0	1	1	0.94	0	8
4	0.99	0.97	1	0	1	0.82	4	3
5	1	0.99	0	0	1	0.88	0	0
6	0.99	0.99	1	0	1	0.95	0	4
7	1	1	0	0	1	0.94	0	1
8	1	1	0	0	1	0.93	0	0
9	0.99	0.98	0	1	1	0.95	0	0
10	1	1	0	0	1	0.97	0	7

R, the level of grey of the circular band is exploited to navigate at roughly constant angular speed around the light. However, in this condition, the amount of light that impinges on the robot's sensor is an important cue for the robot to adjust its discrimination strategies to the characteristics of the environment. It seems that the light is exploited by adaptive regulatory mechanisms that integrate over time the robot's flow of perception, as can be observed by looking at figure 6. This figure shows the
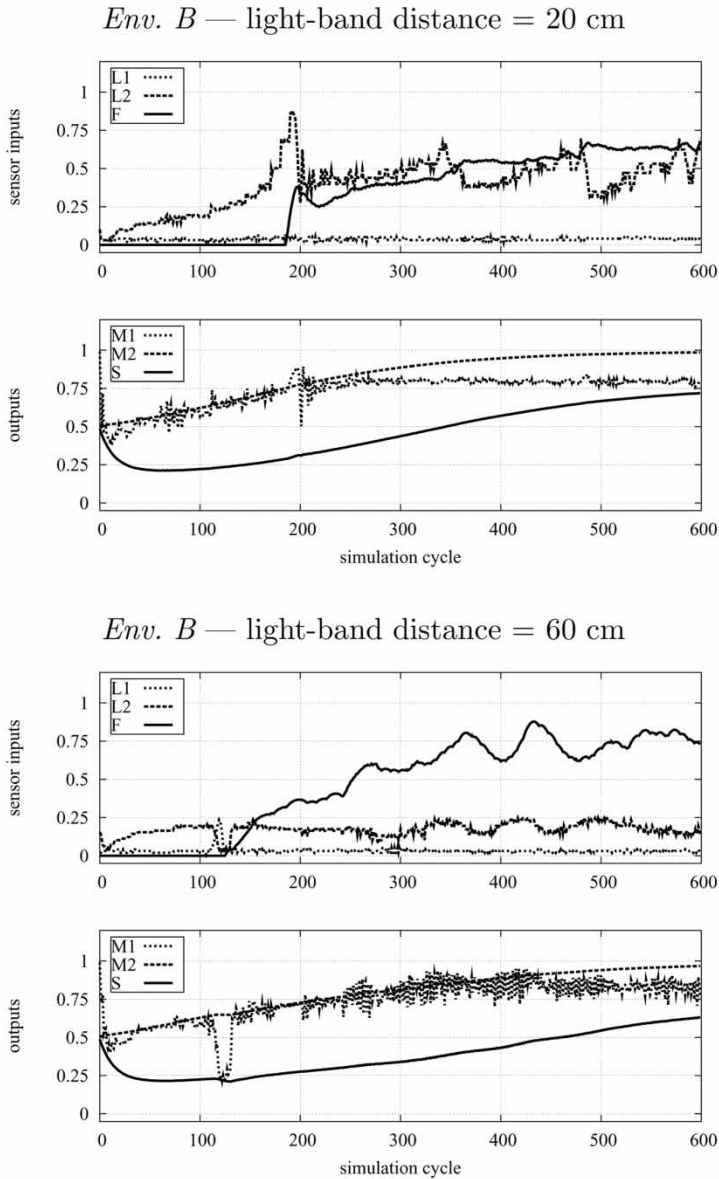

Figure 6. Behavioural analysis. The sensor's activity and the corresponding neural outputs are plotted for 600 simulation cycles. The upper plot refers to Env. B where the light–band distance is 20 cm, while the lower part refers to Env. B having a light–band distance of 60 cm. See the caption of figure 4 for more details about the plots.

sensor's reading, the signal sent by the controller to the robot's wheel, and the normalized activation value of neuron S, i.e. the neuron associated with the sound signalling, of robot run number 10, during a re-evaluation trial in two Env. B: one in which the light–band distance is set to 20 cm (see figure 6 at the top), and one in which the light–band distance is set to 60 cm (see figure 6 at the bottom). From figure 6 it is clear that the activation of the neuron S increases linearly in both environments. However, if we look at the angle of inclination of the continuous line S, we notice that it is steeper in the environment in which the light–band distance is set to 20 cm than in the environment in which the light–band distance is set to 60 cm. Obviously, for a given position of the robot over the circular band, the smaller the light–band distance the higher the light intensity. The robot seems to exploit this feature to adjust its strategy according to the following simple adaptive mechanism: the higher the light intensity, the shorter the time spent by the robot over the circular band required by the neuron S to reach the threshold beyond which the sound is activated.

As we said earlier, adaptive mechanisms evolved in varying environmental circumstances might turn out to be extremely advantageous for achieving a goal other than the one for which they were originally evolved. We have tested, for all the best evolved robots, the robustness of the adaptive mechanisms illustrated above in environmental circumstances never encountered by the robots' ancestors. In particular, we performed an analysis of the robustness of the robot discrimination strategies by looking at the quality of the signalling behaviour, i.e. the offset Δ, in Env. B in which the light–band distance was varying uniformly from 20 to 80 cm. Although in this test several types of environment have never been experienced by the robot's ancestor (i.e. 36 out of 61 environments), we are particularly interested in those in which the light–band distance is greater than 65 cm, because they clearly require generalization capabilities that might not be necessary in the other cases. The offset Δ has been computed 100 times in each environment. The results for the robot run number 10 are shown in figure 7.
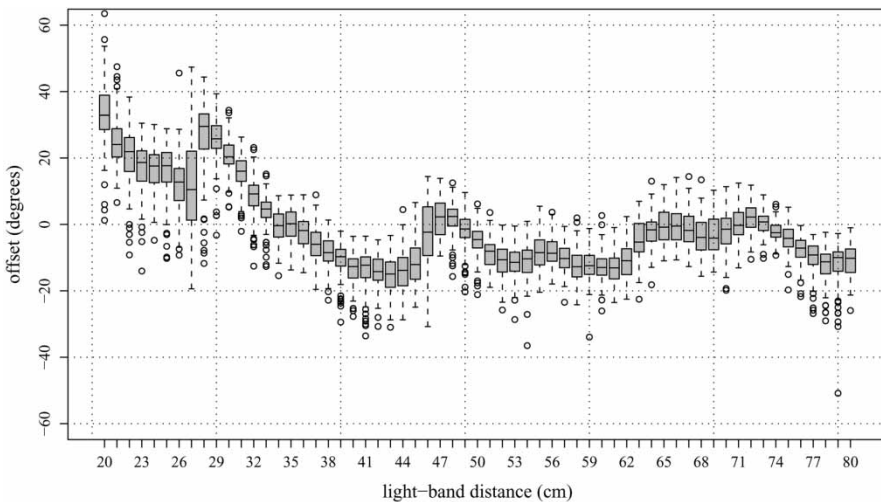

Figure 7. Robustness analysis. The offset Δ is plotted for varying light–band distance. The box-plot shows 100 evaluations per box. Boxes represent the inter-quartile range of the data, while the horizontal bars inside the boxes mark the median values. The whiskers extends to the most extreme data points within 1.5 of the inter-quartile range from the box. The open circles mark the outliers.

Surprisingly, the offset Δ remains within the interval $[-20,40]$ degrees. It takes positive values mainly for small light–band distances. That is, in these environments, the robot tends to signal shortly after having completed a loop around the light. As the light–band distance increases, we can see a progressive decrease in the offset Δ. For light–band distances higher than 29 cm, the offset Δ permanently remains in the interval $[-20,20]$. The precision of the signalling mechanism is not altered by environmental circumstances that have never been encountered by the robot's ancestors (in figure 7, note in particular the performance for light–band distance greater than 65 cm). To conclude, we can say that the discrimination strategies of the best evolved robots proved robust enough to cope with varying environmental circumstances never encountered during evolution. The generalization arises on the one hand because the environment affords the agent to discriminate by employing mechanisms whose functions are not disrupted by the environmental changes; on the other hand, because evolution was able to synthesize time-dependent structures and adaptive mechanisms tailor-made for the agent to capture those 'affordances'.

7. Conclusions

In this paper, we have shown that a single (i.e. not modularized) CTRNN can be shaped by evolution to allow an autonomous agent to make co-ordinated movements that bring forth the perceptual experience necessary to discriminate between two types of environment. The results illustrated here are of particular interest because, contrary to other previous similar studies, in this work the decision-making is uniquely controlled by the time-dependent structures of the agent controller, which in turn are tightly linked to the mechanisms for sensorimotor co-ordination (see section 3).

The first set of simulations, described in section 5, are more focused on the general problem concerning the progressive evolution of behavioural capabilities that are somehow dependent on each other. The results of our simulations show that the evolution of agents capable of solving the considered task is favoured by a particular selective pressure (i.e. condition +A) which facilitates the progressive evolution of fully discriminating agents by firstly rewarding populations of agents capable of sensorimotor co-ordination but not capable of sound signalling, and subsequently by rewarding those agents that combine sensorimotor co-ordination with a proper use of the sound. The other evolutionary scenarios explored, i.e. conditions +B and AB, turned out to be less successful. In particular, condition +B showed that it is far more difficult to associate already evolved sound signalling mechanisms with the mechanisms for the required sensorimotor co-ordination than the other way around. In our case, it was particularly straightforward to isolate the evolutionary dynamics of the condition +B, which hindered the evolution of fully discriminating agents. That is, agents that possessed the mechanisms for signalling and that were not tightly linked with the required sensorimotor co-ordination turned out to be 'lucky' enough to get the highest possible fitness score during the evolutionary phase—the fitness of these agents were clearly overestimated due to the nature of the probabilistic features of the environments in which they happened to be evaluated. The overestimated strategies took over in the population. However, these 'lucky' agents, if re-evaluated in a bigger number of evaluation trials, showed the limitation of their discrimination mechanisms (see table 2).

The results of the second set of simulations, described in section 6, represent, in our view, the most important achievement of our work. These simulations show that, by simply working on the nature of the fitness function, it is possible to bring forth discrimination mechanisms that are robust enough to deal with environmental circumstances

that have never been encountered by the best evolved robot ancestors. Note that these environmental circumstances—that is, the light–band distance—concern the spatio-temporal structures that the robot employs for discrimination (see section 6.2). Therefore, by varying these important environmental structures, we might have induced a particularly disruptive effect on the robot performance. Nevertheless, the robots managed successfully to carry out their task, showing a good performance, as shown in figure 7.

The significance of these results is twofold: on the one hand, they further support the suitability of CTRNNs as controllers for autonomous robots. That is, these results indicate that, despite the complexity of the task, in which mechanisms for sensorimotor co-ordination and for discrimination must be tightly linked, CTRNNs can easily be shaped by evolution to bring forth complex reactive and non-reactive mechanisms within a single non-modularized controller. On the other hand, these results bear upon the significance of the evolutionary approach to robotics. That is, they suggest that the evolutionary approach to robotics is a suitable methodological tool to develop adaptive autonomous agents which, like natural systems, can cope with environmental circumstances never encountered by the agents' ancestors during the evolutionary phase. From an engineering point of view, this is a particularly desirable property to observe in autonomous systems, since it represents a way to overcome successfully the limitations of other more classic approaches to robotics (see Brooks 1991a, b, Harvey *et al.* 1992, Wheeler, 1996, for more on this issue).

Based on these preliminary but encouraging results, in future works, we shall consider more challenging experimental set-ups. In particular, the evolution of time-dependent structures will be associated with functions other than simple discrimination through signalling. Time-dependent structures may be employed to trigger effective alternative activities, as is the case for animal species making decisions about the quality of foraging sites. For example, in an environment with more than one light bulb, the robot might decide to give up circuiting around a light that does not have a *way in* zone, and to move to another light bulb. Another interesting scenario might concern a group of robots engaged in a similar light-approaching task. Let us assume that the band in shades of grey is a trough larger than the diameter of a single robot, and that the robots are capable of self-assembling, as described, for example, in Dorigo *et al.* (2004). In this collective robotics scenario, time-dependent structures and sound signalling can have the function of triggering aggregation and self-assembling of the robots. The robots assembled into a bigger structure might be capable of passing over the trough and reach their goal, i.e. the light bulb.

Community is not responsible for any use that might be made of data appearing in this publication.

Notes

1. The term 'feeling', employed extensively in this document, could be seen as an anthropo morphism. However, in this paper we use it to mean no more than: a robot that manages to discriminate between two types of environment by relying on the persistence over time of a particular perceptual experience.

2. Notice that: (i) the term 'recognition' is used to indicate a sequence of agent's actions produced in response to particular sensors' states; and (ii) the term 'memory' does not necessarily refer to a neural structure internal to the agent's controller.

3. This look-up table is available at http://iridia.ulb.ac.be/ $\sim$ etuci/publications.html

4. Neuron N_1 takes input from the ambient light sensor L_1, N_2 from the ambient light sensor L_2 and N_3 from the floor sensor F.

5. The same experiments performed using a more constraining fitness function yield a success rate of 50% (data not shown).

6. A phylogenetic analysis revealed that the sound signalling behaviour is the last capability to appear among the repertoire of behaviours shown by the evolved robots (data not shown).

7. It is not surprising to observe high error rates in the post-evaluation analysis, even if during the evolutionary phase the best evolved individuals reached the optimal fitness value. This phenomenon is certainly due to an overestimation of the performance of the best-rated individual of the population, which is tested only in a small subset of all the possible environmental circumstances it might encounter.

8. Actually, 3/4 of a loop would be enough, the amplitude of the *way in* zone being fixed to $\pi/2$. This choice has been made for generality purposes.

References

Alphen, J. L. M. V., Bernstein, C., and Driessen, G., 2003, Information acquisition and time allocation in insect parasitoids. *Trends in Ecology and Evolution*, **18**: 81–87.

Beer, R. D., 1995, A dynamical systems perspective on agent–environment interaction. *Artificial Intelligence*, **72**: 173–215.

Blynel, J., and Floreano, D., 2003, Exploring the T-maze: evolving learninglike robot behaviors using CTRNNs. In G. R. Raidl, S. Cagnoni, J. J. R. Cardalda, D. Corne, J. Gottlieb, A. Guillot, E. Hart, C. G. Johnson, E. Marchiori, J.-A. Meyer, and M. Middendorf (eds) *Applications of Evolutionary Computing, EvoWorkshops2003*, Vol. 2611 of *LNCS*, (Berlin: Springer), pp. 598–609.

Bolles, R. C., 1970, Species-specific defence reactions and avoidance learning. *Psychological Review*, **77**: 32–48.

Breland, K., and Breland, M., 1961, The misbehavior of organisms. *American Psychologist*, **16**: 681–684.

Brooks, R., 1991a, Intelligence without representation. *Artificial Intelligence*, **47**: 139–159.

Brooks, R. A., 1991b, Intelligence without reason. In R. Myopoulos, and J. Reiter (eds) *Proceedings of the 12th International Joint Conference on Artificial Intelligence* (San Mateo CA: Morgan Kaufmann), pp. 569–595.

Croon, G. D., Nolfi, S., and Postma, E. O., 2004, *Toward Pro-active Embodied agents: On the Import- ance of Neural Mechanisms Suitable to Process Information in Time* (Complex Engineering Systems, Perseus Books Groups Press) (in press).

Dorigo, M., Trianni, V., Şahin, E., Groß, R., Labella, T. H., Baldassarre, G., Nolfi, S., Deneubourg, J.-L., Mondada, F., Floreano, D., and Gambardella, L. M., 2004, Evolving self-organizing behaviors for a swarm-bot. *Autonomous Robots*, **17**: 223–245.

Goldberg, D. E., 1989, *Genetic Algorithms in Search, Optimization and Machine Learning* (Reading MA: Addison-Wesley).

Gould, S. J., and Vrba, E. S., 1982, Exaptation—a missing term in the science of form. *Paleobiology*, **8**: 4–15.

Harvey, I., Husbands, P., and Cliff, D., 1992, Issues in evolutionary robotics. In J.-A. Meyer, H. Roitblat, and S. Wilson (eds), *From Animals to Animats II: Proceedings of the 2nd International Conference on Simulation of Adaptive Behavior* (Cambridge MA: MIT Press), pp. 364–373.

Jakobi, N., 1997, Evolutionary robotics and the radical envelope of noise hypothesis. *Adaptive Behavior*, **6**: 325–368.

Johnston, T. D., 1981, Contrasting approaches to a theory of learning. *The Behavioral and Brain Sciences*, **4**: 125–173.

Nolfi, S., 2000, *EvoRob 1.1 User Manual*. Institute of Psychology, National Research Council (CNR), http://gral.ip.rm.cnr.it/evorobot/simulator.html

Nolfi, S., 2002, Evolving robots able to self-localize in the environment: the importance of viewing cognition as the result of processes occurring at different time scales. *Connection Science*, **14**: 231–244.

Nolfi, S., and Floreano, D., 2000, *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-organizing Machines* (Cambridge MA: MIT Press).

Nonacs, P., 2003, State dependent behavior and the marginal value theorem. *Behavioral Ecology*, **12**: 71–83.

Parisi, D., Cecconi, F., and Nolfi, S., 1990, Econet: neural networks that learn in an environment. *Network*, **1**: 149–168.

Todd, P. M., and Miller, G. F., 1991a, Exploring adaptive agency II: simulating the evolution of associative learning. In J.-A. Meyer, and S. Wilson, (eds), *From Animals to Animats I: Proceedings of the 1st International Conference on Simulation of Adaptive Behavior* (Cambridge MA: MIT Press), pp. 306–315.

Todd, P. M., and Miller, G. F., 1991b, Exploring adaptive agency III: simulating the evolution of habituation and sensitization. In H. P. Schwefel, and R. Männer (eds) *Proceedings of the 1st International Conference on Parallel Problem Solving from Nature (PPSN)*, Vol. 496 of LNCS, (Berlin: Springer), pp. 307–313.

Trianni, V., Tuci, E., and Dorigo, M., 2004, Evolving functional self-assembling in a swarm of autonomous robots. In S. Schaal, A. Ijspeert, A. Billard, S. Vijayakumar, J. Hallam, and J.-A. Meyer, (eds) *From Animals to Animats VIII. Proceedings of the 8th International Conference on Simulation of Adaptive Behavior* (Cambridge MA: MIT Press), pp. 405–414.

Tuci, E., Harvey, I., and Todd, P. M., 2002a, Using a net to catch a mate: evolving CTRNNs for the Dowry problem. In B. Hallam, D. Floreano, J. Hallam, G. Hayes, and J.-A. Meyer (eds) *From Animals To Animats VII: Proceedings of the 7th International Conference on the Simulation of Adaptive Behavior* (Cambridge MA: MIT Press), pp. 292–302.

Tuci, E., Quinn, M., and Harvey, I., 2002b, An evolutionary ecological approach to the study of learning behaviour using robot-based model. *Adaptive Behavior*, **10**: 201–221.

Wheeler, M., 1996, From robots to Rothko: the bringing forth of world. In M. A. Boden, (ed) *The Philosophy of Artificial Life* (Oxford: Oxford University Press), Chapter 7, pp. 209–236.

Yamauchi, B. M., and Beer, R. D., 1994, Sequential behavior and learning in evolved dynamical neural networks. *Adaptive Behavior*, **2**: 219–246.

Ziemke, T., and Thieme, M., 2002, Neuromodulation of reactive sensorimotor mappings as a short-term memory mechanism in delayed response tasks. *Adaptive Behavior*, **10**: 185–199.