Evolving Neural Mechanisms for an Iterated Discrimination Task: A Robot Based Model

Elio Tuci, Christos Ampatzis, and Marco Dorigo

IRIDIA, Université Libre de Bruxelles - Bruxelles - Belgium
{etuci, campatzi, mdorigo}@ulb.ac.be

Abstract. This paper is about the design of an artificial neural network to control an autonomous robot that is required to iteratively solve a discrimination task based on time-dependent structures. The "decision making" aspect demands the robot "to decide", during a sequence of trials, whether or not the type of environment it encounters allows it to reach a light bulb located at the centre of a simulated world. Contrary to other similar studies, in this work the robot employs environmental structures to iteratively make its choice, without previous experience disrupting the functionality of its decision-making mechanisms.

1 Introduction

Evolutionary Robotics (ER) is a methodological tool for the design of robots' controllers. Owing to its properties, ER can also be employed to study the evolution of behaviour and cognition from a perspective complementary to classic biological/psychological methods (see [4]).

Given the current "status" of their research field, ER practitioners focus not only on studies with an explicit bearing on engineering or biological literature but also on studies which aim to further develop their methods. For example, several research works have focused on the modelling and exploitation of alternative controllers for autonomous robots—e.g., spiking networks [7], and gas networks [5]. In general, these works look at how to exploit evolution to shape these controllers rather than at the complexity and the significance of the evolved behaviour. Contrary to these, other works are more focused on the evolution of novel—i.e., never evolved yet—and complex behaviour. For example, some works exploited "classic" neural structures to evolve controllers for agents capable of non-reactive or learning behaviour [8]. The results of these studies should be considered as a "proof-of-concept": they show that the type of control structure employed can be shaped by evolutionary algorithms to provide the robot with the underlying mechanisms required to solve the task at hand.

The work illustrated in this paper belongs to the latter category. To the best of our knowledge, this is the first study in which a single (i.e., non modularised) dynamic neural network has been shaped to control the behaviour of an autonomous robot engaged in an iterated discrimination task.[1]. The task requires

[1] A literature review of the field can be found in [9].

M. Capcarrere et al. (Eds.): ECAL 2005, LNAI 3630, pp. 231–240, 2005.

navigation within a circular arena in order to approach a light bulb located at the centre of this simulated world. The "decision making" aspect requires the robot "to iteratively decide", during a sequence of trials, whether or not the types of environment it encounters allow it to accomplish its task. The difficulty of the task lies, on the one hand, in the nature of the discrimination problem, which requires the integration of sensory information over time; on the other hand, in the design of decision making mechanisms to carry out the iterated discrimination task. That is, this task requires the robot to possess memory structures which do not lose their functionality due to potentially disruptive effects of the previous experience—i.e., the nature and the amount of discriminations already made. The results show that dynamic neural networks can be successfully designed by evolution to allow a robot to iteratively solve the discrimination task based on time-dependent cues. We also provide an analysis which gives some hints on the strategy employed by the best evolved robot to solve the task.

2 Methods

Description of the task. At the start of each trial, the simulated robot is placed in a circular arena with a radius of 110 cm (see Fig. 1). The arena is simulated as a toroidal world; that is, if the robot traverses the world's boundaries from one side, it comes in from the other side at the anti-diametrical position. At the centre of this world there is a light bulb that is always turned on during a trial. The light can be perceived up to a distance of 90 cm. Between 90 cm and 110 cm of distance to the bulb, the robot does not perceive any light. We refer to this area of the arena as the *dark* zone. The robot perceives the light through its ambient light sensors. 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 cm and 60 cm from the light. The band is divided in three sub-zones of equal width but coloured differently—i.e., light

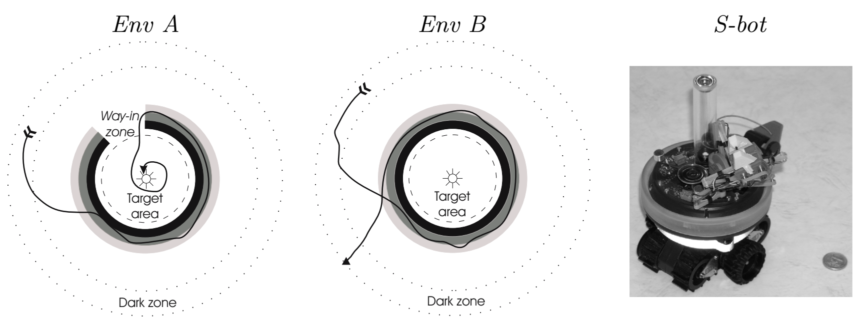

Fig. 1. Depiction of the task and picture of a *s-bot* close to a 1 Euro coin. The *dark* zone is the area within the dotted circles. The target area, centred on the light, is indicated by the dashed circle. The continuous arrows are examples of good navigational strategies.

grey, dark grey, and black. The robot perceives the colour of the floor through its floor sensor, positioned under its chassis.

The robot can freely move within the arena as well as on the circular band, but it is not allowed to cross the black edge of the band close to the light. This edge can be imagined as an obstacle or a trough, that prevents the robot from further approaching the light. Whenever the robot crosses the black edge, the trial is unsuccessfully terminated. The light grey and the dark grey zones are 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 Fig. 1, 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 Fig. 1, middle). The *way in* zone represents the path along which the robot is allowed to safely reach the light in *Env A*. The robot cannot reach the proximity of the light in *Env B*.

At the start of each trial, the robot does not know in which type of environment it is located. It finds itself positioned in the *dark* zone with a random orientation. At this time its task is to explore the arena, in order to get as close as possible to the light. If it encounters the circular band it has to start looking for the *way in* zone in order to continue approaching the light. If it finds the *way in* zone, the robot has to get closer to the light and remain in its proximity for 10s. After this time, the trial is successfully terminated and the robot is randomly re-positioned in the *dark* zone. If there is no *way in* zone (i.e., the current environment is *Env B*), the robot should be capable of (a) "recognising" the absence of the *way in* zone, (b) notifying by a sound signal the absence of the *way in* zone, (c) coming back to the *dark* zone by performing anti-phototaxis. Back in the *dark* zone either because re-positioned or because returned there, the robot has to "prepare" itself for a new trial in which the characteristics of the environment are unknown. The transition between two consecutive trials is particularly complex in case the robot has just concluded a trial in *Env B*. This transition requires the robot to turn the sound off and to switch from anti-phototaxis (i.e., the last behaviour performed in *Env B*) to random walk and then phototaxis once the light falls again within its perceptual field.

This task is very similar to the one described in [9] since the robot is required to make use of a temporal cue in order to discriminate between *Env A* and *Env B*. This discrimination is based on the persistence of the perception of a particular sensorial state (e.g., the perception of the grey floor, the light, or both) for the amount of time that, given the trajectory and speed of the robot, corresponds to the time required to make a loop around the light. In other words, if the perception of a particular sensorial state common to both types of environment lasts significantly long with respect to the speed and trajectory of the robot, then that sensorial state might be used by the robot to "conclude" that there is no *way in* zone, and a tone has to be emitted (see [9] for more details).

However, with respect to [9], this task is meant to be a step further in the evolution of decision making mechanisms based on time-dependent structures.

In [9], we studied the evolution of decision making mechanisms for a one shot discrimination task by simply resetting the robot's controller (i.e., set to 0 the cell potential of the neurons) at the beginning of each trial. The resetting "facilitates" the task of discriminating between *Env A* and *Env B* since (a) the integration of the sensorial state which eventually leads to the emission of the sound signal is not disrupted by the type and the amount of previous experience; (b) the robot does not need to terminate the emission of the sound signal, since, given the way in which sound is implemented, such an event is automatically determined by the resetting; (c) the robot does not need to "recognise" the end of the current trial and the beginning of a new one, since such transition implies the resetting of the activation values of the neurons of its controller. In other words, each trial is for the robot like a new life in which, starting from the same internal state, a single decision has to be made.

The task described in this paper is made significantly more complex with respect to what shown in [9] by (a) avoiding to impose the resetting of the robot controller at the beginning of each trial, and consequently by (b) letting the robot autonomously develop the conditions which set the end of a trial and the beginning of a new one. If the robot controller is not reset at the beginning of a trial, the decision to be made in the trials following the first one, will necessarily be carried out by mechanisms which have already been "shaped" by previous experience.[2] Therefore, it is important that the functionality of the decision making mechanisms employed by the robot are not disrupted by previous experience. In other words, discriminating between *Env A* and *Env B* requires the robot to make use of memory structures to integrate over time a particular sensorial state. Carrying out such an iterated discrimination task requires the robot to possess memory structures which do not lose their functionality due to potential disruptive effects of the previous experience—i.e., the nature and the amount of discriminations already made. Furthermore, the robot should be able to exploit its perception in order to establish when a trial ends and a new one starts. This is particularly important at the end of an exploration in *Env B*, in which the robot should conclude the trial by emitting a tone and moving away from the light and should begin the new trial with the sound turned off and performing light seeking behaviour. These changes (i.e., sound on - sound off, anti-phototaxis - phototaxis) have to be triggered by perceptual states which ideally set the end of a trial and the start of a new one.

Several implementation details such as (i) requiring the robot to perform anti-phototaxis in *Env B*, (ii) the introduction of the *dark* zone, and (iii) the toroidal world, have been introduced to make sure that the robot's sensory experience can potentially provide the support the robot needs in order to make iterated choices. For example, a robot that successfully terminates a trial in *Env B* can exploit the perceptual states associated with performing anti-phototaxis and with its presence in the *dark* zone to "prepare" itself for the new trial (a) by turning the sound signalling off, and (b) by adjusting its internal state so that it will be ready for a new discrimination task. In particular, being repositioned in

[2] In our model, it is the neuron's cell potential to be modified by the robot's experience.

the *dark* zone after a success in *Env A*, or reaching the *dark* zone after a success in *Env B*, are two events that can be unambiguously employed by the robot in order to establish the end of the current trial and the beginning of the following one. In the absence of a global framework of reference (e.g., a compass), the toroidal world makes it easier for a robot to navigate in the *dark* zone in order to reach the area in which the light source can be perceived.

The simulation model. The robot and its world are simulated using the "minimal simulation" technique described in [6]. This technique uses high levels of noise to guarantee that the simulated controller transfers to the physical robot with no loss of performance (see [1]). Our simulation models some of the hardware characteristics of the real *s-bots*. The *s-bots* are small wheeled cylindrical robots, 5.8 cm of radius, equipped with infrared proximity sensors, light and humidity sensors, accelerometers, and omni-directional camera (see Fig. 1, right, and also `http://www.swarm-bots.org/` for more details). In particular, our simulated *s-bot* is provided with four ambient light sensors, placed at $-112.5°$ (A_1), $-67.5°$ (A_2), $67.5°$ (A_3), and $112.5°$ (A_4) with respect to its heading, a floor sensor positioned facing downward on the underside of the robot (F), an omni-directional sound sensor (S), and a loud-speaker. The motion of the robot is implemented by the two wheel actuators. Light levels change as a function of the robot's distance from the lamp. Light sensor activation values are taken from a look-up table which contains sampled information from the real robot. The ground sensor detects the level of grey of the floor. The robot floor sensor outputs the following values: 0 if the robot is positioned over the white floor; $\frac{1}{3}$ if the robot is positioned over the light grey floor; $\frac{2}{3}$ if the robot is positioned over the dark grey floor; 1 if the robot is positioned over the black floor. The simulated speaker produces a binary output (on/off); the sound sensor has no directionality and intensity features. Concerning the function that updates the position of the robot within the environment, we employed the Differential Drive Kinematics equations, as presented in [2]. 10% uniform noise was added to the light sensor readings, the motor outputs and the position of the robot.

The controller and the evolutionary algorithm. Fully connected, eight neuron continuous time recurrent neural networks (CTRNNs) are used. All neurons are governed by the following state equation:

$$\frac{dy_i}{dt} = \frac{1}{\tau_i}\left(-y_i + \sum_{j=1}^{8}\omega_{ji}\sigma(y_j+\beta_j) + gI_i\right) \qquad \sigma(x) = \frac{1}{1+e^{-x}} \tag{1}$$

where, using terms derived from an analogy with real neurons, y_i represents the cell potential, τ_i the decay constant, β_j the bias term, $\sigma(y_j+\beta_j)$ the firing rate, ω_{ji} the strength of the synaptic connection from neuron j^{th} to neuron i^{th}, I_i the intensity of the sensory perturbation on sensory neuron i. Four neurons receive input (I_i) from the robot sensors: neuron N_1 takes input from $\frac{A_1+A_2}{2}$, N_2 from $\frac{A_3+A_4}{2}$, N_3 from F, and N_4 from S. Neurons N_1, N_2, and N_3 receive as input a real value in the range [0,1], while neuron N_4 receives a binary input (i.e., 1 if a tone is emitted, otherwise 0). The other neurons do not receive any input from the robot's sensors. The cell potential (y_i) of the 6^{th} neuron, mapped into [0,1] by a sigmoid function (σ), is used by the robot to control the sound signalling system (i.e., the robot emits a sound if $y_6 >= 0.5$). The cell potentials (y_i) of the 7^{th} and the 8^{th} neuron, mapped into [0,1] by a sigmoid function (σ) and then linearly scaled into $[-6.5, 6.5]$, set the robot motors output. The strength of synaptic connections ω_{ji}, the decay constant τ_i, the bias term β_j, and the gain factor g are genetically encoded parameters. Cell potentials are set to 0 any time the network is initialised or reset, and circuits are integrated using the forward Euler method with an integration step-size of 0.1.

A simple generational genetic algorithm is employed to set the parameters of the networks [3]. The population contains 100 genotypes. Generations following the first one are produced by a combination of selection with elitism, recombination and mutation. For each new generation, the three highest scoring individuals ("the elite") 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, 8 decay constants, 8 bias terms, and a gain factor). Initially, a random population of vectors is generated by initialising each component of each genotype to values chosen uniformly random from the range [0,1]. New genotypes, except "the elite", 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 0.13. The mean of the Gaussian is 0, 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 $\omega_{ji} \in [-6, 6]$ and gain factor $g \in$ [1,12]. Decay constants are firstly linearly mapped onto the range $[-0.7, 1.7]$ and then exponentially mapped into $\tau_i \in [10^{-0.7}, 10^{1.7}]$. The lower bound of τ_i corresponds to a value slightly smaller than the integration step-size used to update the controller; the upper bound corresponds to a value slightly bigger than the average time required by a robot to reach and to perform a complete loop of the band in shades of grey.

The evaluation function. During evolution, each genotype is coded into a robot controller, and is evaluated ten times, 5 trials in *Env A*, and 5 trials in *Env B*. The sequence of environments within the 10 trials is chosen randomly. Each trial (e) differs from the others in the initialisation of the random number generator, which influences the robot starting position and orientation, the position and amplitude of the *way in* zone, and the noise added to motors and sensors. The width of the *way in* zone can vary from 45° to 81°. Within a trial, the robot life-span is 80 s (800 simulation cycles). In each trial, the robot is rewarded by an evaluation function f_e which corresponds to the sum of the following four components:

1) C_1 rewards fast movement to the target area. $C_1 = \frac{d_i - d_c}{d_i}$ where d_i and d_c represent respectively the initial and the current Euclidean distance between the robot and the light bulb. In *Env A*, C_1 is set to 1 if the robot terminates the trial less than 35 cm away from the light bulb. In *Env B*, C_1 is set to 1 as soon as the robot reaches the circular band without crossing the black edge in the direction of the light.
2) C_2 rewards movements away from the light. $C_2 = \frac{d_c}{d_{max}}$ if trial in *Env B*, 0 if trial in *Env A* or if $C_1 < 1$ ($d_{max} = 110$ cm).
3) C_3 rewards agents that never signal in *Env A* and that always signal in *Env B*. C_3 is set to 1 if the robot signals properly, 0 otherwise. The robot is considered to have signalled only if it has done so being closer than 70 cm from the light. By doing so, we create an area between 70 cm and 110 cm from the light that the robot can use to turn the sound off at the end of a trial in *Env B*.
4) C_4 rewards movements toward the light. $C_4 = 1 - \frac{k}{T}$ if trial in *Env A*, 0 otherwise. k is the number of simulated time-steps the robot spent to reach the target area, and $T = 800$ is the total number of simulated time-steps available to the robot. 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.

3 Results

Ten evolutionary simulations, each using a different random initialisation, were run for 7000 generations. We examined the best individual of the final generation from each of these runs in order to establish whether they evolved the required behaviour. Recall that the robot is successful in *Env A* if it reaches the target area without emitting any sound signal; it is successful in *Env B* if (a) it reaches the circular band, (b) signals the absence of the *way in* zone by emitting a tone, and (c) comes back to the *dark* zone (anti-phototaxis).

During the post-evaluation phase, each of the ten best evolved controllers was subjected to a set of 252 different re-evaluations. Since a re-evaluation is composed of 10 trials, out of which 5 are *Env A* and 5 are *Env B*, 252 ($\frac{10!}{5! \cdot 5!}$) is the number of possible evaluations which differ in the order of presentation of *Env A* and *Env B*. 2520 is the total number of trials experienced by each robot during the post-evaluation, half of which in *Env A* and half in *Env B*. Note that during evolution, each robot experienced only a particular sequence of 5 trials in *Env A* and 5 trials in *Env B*. Since the robot controller is reset only at the beginning of each evaluation, the order of presentation of the types of environment might bear upon its performance. A robot that results successful in the post-evaluation is one which employs a strategy which is effective regardless the sequence of environments.

During the post-evaluation phase, we looked at the robot's capability to reach the light bulb (*Succ.*) in *Env A*, without making any error. Errors can be of three types: error (I) refers to the case in which the robot emits a sound signal, error (II) refers to the case in which the robot crosses the black edge of the band, error

Table 1. Results of post-evaluation showing the performance of the best evolved controllers of each run. The percentage of success (*Succ.*) and the percentage of errors (I, II, III in *Env A*, and IV, V, VI, VII in *Env B*) over 252 evaluations are shown for both *Env A* and *Env B*. Additionally, the average offset Δ, its standard deviation (degrees), and the number of successful trials (*n.*) are shown for *Env B*.

run	*Env A*				*Env B*							
	Succ.	*Types of Error (%)*			*Succ.*	*Types of Error (%)*				offset Δ		
		I	II	III		IV	V	VI	VII	*Avg.*	*Std*	*n.*
n. 1	83.17	7.77	4.76	0.15	1.42	14.92	0.00	4.12	79.52	21.25	115.86	18
n. 2	94.12	5.87	0.00	0.00	96.42	1.34	0.00	0.00	2.22	50.52	124.08	1215
n. 3	22.85	0.00	77.06	0.0	0.0	0.23	0.0	99.76	0.0	—	—	0.0
n. 4	84.92	14.68	0.23	0.00	98.57	0.0	0.0	0.0	1.42	-81.12	39.32	1242
n. 5	86.19	8.33	5.23	0.07	81.58	0.07	0.15	15.23	2.93	-10.8	79.05	1028
n. 6	33.17	10.07	50.55	6.19	92.93	0.00	0.00	0.079	6.98	6.1	104.75	1171
n. 7	88.49	11.11	0.39	0.00	97.53	0.00	0.00	0.00	2.46	-15.81	80.87	1229
n. 8	82.93	16.50	0.47	0.07	96.19	1.11	0.00	0.00	2.69	-60.04	82.16	1212
n. 9	99.68	0.07	0.23	0.00	95.87	2.06	0.00	0.23	1.82	71.03	50.89	1208
n. 10	59.04	40.95	0.00	0.00	98.57	0.07	0.00	0.00	1.34	-155.86	57.36	1242

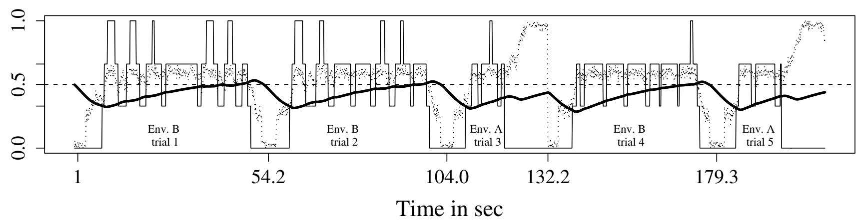

Fig. 2. The graph shows the output of the neuron that controls the sound (N_6, see continuous thick line), the floor sensor reading (F, see continuous thin line), the average values of the light sensor readings (A_1, A_2, A_3, A_4, see dotted line), during the first 5 trials of a 10 trials evaluation in which the robot did not make any error. The numbers on the x-axis, show at which point of the robot life-span a new trial begins.

(III) refers to the case in which the robot makes both errors I and II within the same trial. Similarly, in *Env B*, we looked at the performance of the robot in completing the task as mentioned above (*Succ.*), without committing any error. Four error types are possible: error (IV) refers to the lack of sound signalling, error (V) refers to the robot crossing the black edge of the band, error (VI) refers to the robot missing to reach the *dark* zone after having signalled; error (VII) refers to the case in which the robot makes error IV, V, and VI within the same trial. Furthermore, in *Env B* we also compute the offset (offset Δ) between the entrance position of the robot in the circular band and the position in which the robot starts to signal (see [9] for a description of how the offset Δ is computed). This measure accounts for the precision of signalling with respect to the time it takes for the robot to complete a loop around the light. Offset Δ takes value 0° if the robot signals exactly after covering a complete loop of the circular band. Negative values of the offset Δ suggest that the robot signals before having performed a complete loop; positive values correspond to the situation in which the robot emits a tone after having performed a loop around the light.

The results of the post-evaluation, shown in Table 1, shed light on two aspects of our work: first, they give a quantitative estimate of the overall success rate of the evolved strategies; second, they provide elements to infer the behavioural strategies employed by our robots to solve the task. Concerning the percentage of success in both types of environment, the results are quite encouraging. Despite the complexity of the task, six runs out of ten—runs n. 2, 4, 5, 7, 8, 9—show a percentage of success (*Succ.*) in both types of environment higher than 80% (i.e., more than 2016 successful trials out of 2520). The strategies of run n. 2 and 9 are the most effective, with a percentage of success in both environments higher than 94%. The performance of run n. 4, 5, 7, 8 is mainly "blurred" by errors of type I, caused by a risk-taking behaviour, which led the robot to signal slightly before having completed a loop around a light—see the negative values of the average offset Δ. Among the less successful robots, the performance of run n. 10 is also disrupted by errors of type I. The bad result of run n. 6, and n. 2 is mainly due to crossing the black edge of the circular band (error type II). Run n. 1 is quite successful in *Env A*, but its performance is particularly bad in *Env B*. In view of its high error rate of type VII, we can conclude that this robot employs the strategy of never signalling in *Env B*, and of remaining on the circular band circuiting around the light. It is worth noticing that the two most successful runs (i.e., run n. 2 and 9) employ a risk-averse behaviour, since they have the tendency to signal slightly after having completed a loop around a light—see the positive values of the average offset Δ.

The graphs shown in Fig. 2 give us some hints on the mechanisms employed by robot run n. 9 (a) to control the sound, and (b) to switch from phototaxis to anti-phototaxis and vice-versa. As far as it concerns (a), Fig. 2 shows that the output of neuron N_6 increases if the robot is on the circular band. The output of N_6 crosses the 0.5 threshold—i.e., the sound is turned on—if the robot remains on the band for a time slightly higher than the time required to make a loop around the light. This can be inferred by the positive value of the offset Δ in Table 1. Concerning (b), the robot performs phototaxis as long as it does not perceive any sound. Fig. 2 shows that the perception of sound makes the robot change strategy—i.e., from phototaxis to anti-phototaxis. Once the robot is out of the band, the output of neuron N_6 starts decreasing. However, given the rate of change of the output of neuron N_6, the robot stops emitting a tone just after having reached the *dark* zone. Owing to this mechanism, the robot manages, by the end of a trial in *Env B*, to set the cell potential of neuron N_6 to a value which makes it possible for it (a) to approach the light at the beginning of the following trial, since no sound is emitted, and (b) to be in a "state" to be able to perform another discrimination.

4 Conclusions

In this paper, we have shown that a single dynamic neural network can be synthesised by evolution to allow an autonomous robot to successfully perform an iterated discrimination task, based on time-dependent structures. The results illustrated here are of particular interest because, to the best of our knowledge, this is the first study in which an autonomous robot manages to iteratively solve a complex non-reactive task without previous experience disrupting its decision making mechanisms. The performance of the best evolved robot was not disrupted by the sequence of presentation of the environments. However, the robustness of the evolved strategies with respect to other potentially more disruptive environmental changes, such as the dimensions of the circular band, and the *dark* zone, remains to be assessed.

Acknowledgements. This work was supported by the *ECAgents* project, funded by the Future and Emerging Technologies programme of the European Commission, under grant IST-1940. The authors thank Thomas Halva Labella and Vito Trianni for stimulating discussions and feedback during the preparation of this paper, and the three reviewers for their helpful comments. Marco Dorigo acknowledges support from the Belgian FNRS, of which he is a Research Director, and from the "ANTS" project, an "Action de Recherche Concertée" funded by the Scientific Research Directorate of the French Community of Belgium. The information provided is the sole responsibility of the authors and does not reflect the Community's opinion. The Community is not responsible for any use that might be made of data appearing in this publication.

References

1. C. Ampatzis, E. Tuci, V. Trianni, and M. Dorigo. Evolving communicating agents that integrate information over time: a real robot experiment. Technical Report TR/IRIDIA/2005-12, Université Libre de Bruxelles, 2005.
2. G. Dudek and M. Jenkin. *Computational Principles of Mobile Robotics*. Cambridge University Press, Cambridge, UK, 2000.
3. D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA, 1989.
4. I. Harvey, E. A. Di Paolo, R. Wood, M. Quinn, and E. Tuci. Evolutionary robotics: A new scientific tool for studying cognition. *Artificial Life*, 11(1–2):79–98, 2005.
5. P. Husbands, T. Smith, N. Jakobi, and M. O'Shea. Better living through chemistry: Evolving GasNets for robot control. *Connection Science*, 10(3–4):185–210, 1998.
6. N. Jakobi. Evolutionary robotics and the radical envelope of noise hypothesis. *Adaptive Behavior*, 6:325–368, 1997.
7. E. A. Di Paolo. Evolving spike-timing dependent plasticity for single-trial learning in robots. *Philosophical Transactions of the Royal Society A*, 361:2299 –2319, 2003.
8. E. Tuci, M. Quinn, and I. Harvey. An evolutionary ecological approach to the study of learning behaviour using a robot-based model. *Adaptive Behavior*, 10(3-4):201–221, 2003.
9. E. Tuci, V. Trianni, and M. Dorigo. 'Feeling' the flow of time through sensory/motor coordination. *Connection Science*, 16:1–24, 2004.