

# Evolving communicating agents that integrate information over time: a real robot experiment

Christos Ampatzis, Elio Tuci, Vito Trianni and Marco Dorigo

IRIDIA - Université Libre de Bruxelles, Bruxelles, Belgium  
{campatzi,etuci,vtrianni,mdorigo}@ulb.ac.be,  
<http://iridia.ulb.ac.be/>

**Abstract.** In this paper we aim to design an artificial neural network to control a homogeneous group of two autonomous robots that are required to solve a discrimination task based on time-dependent structures. The network should produce alternative actions according to the discrimination performed. We evolved controllers that have been successfully ported on the real robots. We also show that the system displays a simple form of communication among the agents.

## 1 Introduction

In recent years, Evolutionary Robotics (ER) has been used as a method both to engineer control structures for autonomous robots, and to study the evolution of behaviour from a perspective complementary to biological analytical models [9]. As an engineering method, ER offers the possibility of automating a complex design task. 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 design space that conventional design approaches are often constrained to ignore (e.g., see [11]). As a biological tool, ER offers the possibility to look at the effects of the physical interactions between an embodied and situated agent and its environment, on the evolution of behaviour (e.g., see [12]).

Due to the number of trials needed to test individuals, the design of robot controllers by means of artificial evolution is generally carried out by using simulation models. The latter are digital media that reproduce some of the physical phenomena that can be observed when a real robot is located and acts in a real world [5,7]. By speeding up the evaluation time, the use of simulation models allows the ER practitioners to save time in the control-policies design process. However, as claimed in [1], “There is real danger that programs which work well on simulated robots will completely fail on real robots ...”. The digital medium might miss to take into account phenomena that bear upon the functional properties of the evolved controllers. As a consequence, the latter result less effective in managing the real world sensing and actuation (see also [6]).

The simulation-reality gap might be harder to cross whenever the robot’s neural mechanisms are “sensitive” to a broad range of spatial and temporal phenomena (i.e. how the sensor readings unfold in time). Noise and unpredictable events inevitably make the dynamics of the real world somehow different from those experienced in simulation [1]. These differences might not change the way the robot reacts at any time cycle to its sensor readings, nevertheless, in a non-reactive task, the same phenomena might severely disrupt the robot’s memory structures and its capability to bring forth adaptive behaviour such as learning.

In spite of this, the recent advances in the design of artificial recurrent neural networks controllers for robots engaged in non-reactive tasks have been hardly validated by using real robots (e.g., see [14,13,2,10]). We believe that, unless the “porting” is experimentally proved to be successful, it is always questionable whether non-reactive mechanisms for memory and learning evolved in a simulated world can deal with reality.

Within the context of the ECAgents project<sup>1</sup>, we intend to carry out research work focused on the design of dynamic neural networks to control group of homogeneous robots that use communication to “inform” each other about things that are not currently accessible to the whole group: e.g., a robot’s own experience and knowledge of parts of the environment not explored yet by other members of the group. These studies are motivated by the following reasons: (i) from an engineering point of view, we are interested in exploiting the evolution of communicative behaviour to optimise the efficiency of the group in tasks in which a “collective intelligence” may be beneficial; (ii) from a biological perspective, these kind of scenarios might be used as robot-based models to investigate issues concerning the evolution of communication and foraging group strategies in natural organisms.

Since the design of non-reactive neural mechanism through evolution requires many evaluation trials, our studies strongly rely on a simulation environment (see Section 2). The aim of this paper is to show the results of a first set of experiments in which non-reactive controllers for learning and communicative behaviour have been initially evolved in a simulated environment and subsequently ported on real robots. We believe that going through the process of porting the controller on the real robot is a necessary step to validate the results of our research work. Section 3 shows a series of analysis which illustrate how different are the performances of the real robots with respect to the simulated ones. Conclusions are drawn in Section 4.

## 2 Methods

In this section, we describe the setup of the evolutionary experiments we conducted. The goal of these experiments is to show how a group of robots can perform a discrimination task between two different environments, based only on temporal cues (i.e., integration over time of the robots’ perceptions) and/or on communication signals. In the following, we describe in detail the task, the simulation model, the controller and evolutionary algorithm and the evaluation function used.

**Description of the task** At the start of each trial, two simulated robots are placed in a circular arena with a radius of 110 cm (see Fig. 1), at the centre of which a light bulb is always turned on. The robots perceive the light through their ambient light sensors. The colour of the arena floor is white except for a circular band, centred around the lamp that 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 grey, dark grey, and black. Each robot perceives the colour of the floor through its floor sensors, positioned under its chassis.

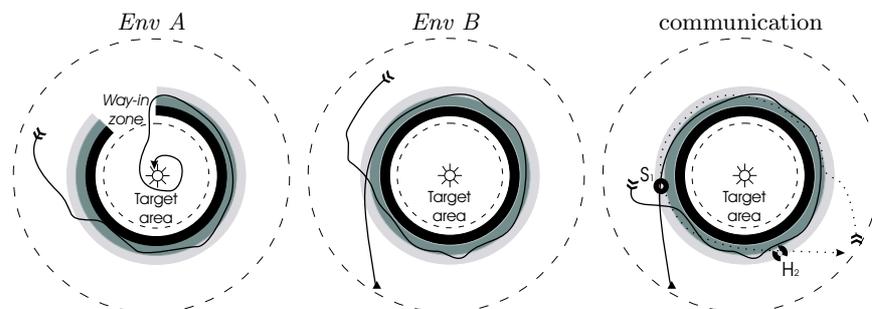
---

<sup>1</sup> ECAgents is a project sponsored by the Future and Emerging Technologies program of the European Community (IST-1940), <http://ecagents.istc.cnr.it/>

The primary goal of the robots is to approach the light bulb. They can freely move within the arena as well as on the circular band, but they are 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 a robot crosses the black edge, the trial is unsuccessfully terminated for that robot. The light grey and the dark grey zones are meant to work as a warning signal which indicate to each 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, called the *way in zone*, where the floor is white (see Fig. 1, left). In the other type, referred to as *Env B*, the band completely surrounds the light (see Fig. 1, center). The *way in zone* represents the path along which the robots are allowed to safely reach the light in *Env A*. On the contrary, they cannot reach the proximity of the light in *Env B*, and in this situation their goal is to leave the band.

At the beginning of each trial, the robots do not know in which type of environment they are located. They are initialized at a distance between 75 and 95 cm from the light, with a random orientation between  $-120^\circ$  and  $+120^\circ$  with respect to the light. At this moment the task of the robots is to explore the arena, in order to get as close as possible to the light. If they encounter the circular band they have to start looking for the *way in zone* in order to continue approaching the light, and once they find it, they should get closer to the light and remain both in its proximity for 30 sec. After this time interval, the trial is successfully terminated. If there is no *way in zone* (i.e., the current environment is an *Env B*), the robots 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) leave the band by performing anti-phototaxis.

Robots are provided with a dedicated communication channel, so whenever a robot produces a signal (“talker”), this signal is “heard” by the other one. Due to this channel, the system can display a simple form of co-operation among its members. In detail, once a robot signals the absence of a *way in zone* in *Env B*,



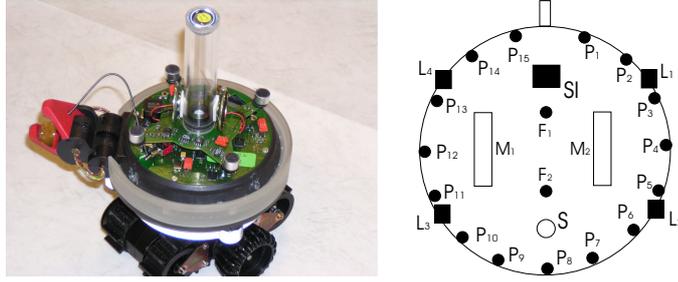
**Fig. 1.** The task. *Env A* (left) is characterised by the *way in zone*. The target area, centred on the light, is indicated by the dashed circle. In *Env B* (center) there is no *way in zone* and the target area cannot be reached. The continuous arrows are an example of a good navigational strategy for one robot. Right: the influence of communication in *Env B*: one robot signals earlier ( $S_1$ ), the other listens ( $H_2$ ) and they both leave the band. The trajectories are displayed with a continuous line for the first robot and a dashed line for the second one.

the other robot can “listen” to this signal and leave the band even though it has not yet gathered enough information in order to conclude in which environment it is located (see Fig. 1, right). Of course, it is up to the evolutionary process to produce agents that *react* to the signal of the other robot and accordingly modify their behaviour or simply *ignore* it and carry on with their task without alterations. In the former case we can say that communication has evolved, while in the latter not.

To summarise, a trial is successful in *Env A* if both robots reach the light without having emitted a signal and without crossing the black edge of the band. In *Env B* though, things are significantly different, since we consider successful a robot that ends up away from the light, therefore at 110 cm from it. This could be the result of either its own “realisation” that there is no *way in* zone, or of the reaction to the signal emitted by the other member of the group. Finally in both environments the robots are not allowed to crash against each other, therefore should prove capable of avoiding each other, displaying obstacle avoidance capabilities. During evolution, the emergence of robot-robot avoidance has been favoured by the following implementation choices: (a) the robots are initialised half of the times at the beginning of each trial close to each other, with a distance between 3 and 20 cm and (b) before a trial is successfully concluded in *Env A*, we force the robots to interact for 30 sec in a very small area of 25 cm around the light, called the “target area” (see Figure 1).

This task is very similar to the one described in [15] since each 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. Nevertheless, this task is made significantly more complex. This time, the integration over time of the robots’ sensorial inputs is used to trigger an alternative action, leaving the band in shades of grey. The same network is responsible of producing all the required actions, depending on the environment in which the robot is placed. Furthermore, by using the robots’ proximity sensors, we allow interactions among the agents that might affect the dynamics of our controller and even interfere with the discrimination task. Finally, we use a homogeneous group of robots, that is the same neural controller is cloned on both robots. Since our task is a collective robotics task, we are interested in the evolution of communication (a robot’s reaction to the signal produced by another robot) and in its effects on the overall performance of the system.

**The simulation model** The robot and its world are simulated using the “minimal simulation” technique described in [5]. This technique uses high levels of noise to guarantee that the simulated controller will transfer to a physically realised robot with no loss of performance. Our simulator models some of the hardware characteristics of the real *s-bots* [8]. The *s-bots* are small wheeled cylindrical robots, 5.8 cm of radius, whose mobility is ensured by a differential drive system composed by tracks and wheels. They are equipped with infrared proximity sensors, light and humidity sensors, accelerometers, omni-directional camera, rotating base and rigid and semi-rigid grippers (see Fig. 2 left, and also <http://www.swarm-bots.org/> for more details). In this work, we make use of four ambient light sensors, placed at  $-112.5^\circ$  ( $L_1$ ),  $-67.5^\circ$  ( $L_2$ ),  $67.5^\circ$  ( $L_3$ ), and  $112.5^\circ$  ( $L_4$ ) with respect to its heading, fifteen infra-red proximity sensors



**Fig. 2.** A picture of an *s-bot* on the left. Plan of the robot on the right, showing sensors and motors. The robot is equipped with four ambient light sensors ( $L_1$  to  $L_4$ ), two floor sensors  $F_1$  and  $F_2$ , 15 proximity sensors ( $P_1$  to  $P_{15}$ ) and a binary sound input sensor, called  $SI$ . The wheel motors are indicated by  $M_1$  and  $M_2$ . A simple sound signalling system is referred to as  $S$ .

placed around its turret ( $P_1$  to  $P_{15}$ ), two floor sensors  $F_1$  and  $F_2$  positioned facing downward on the underside of the robot with a distance of 4.5 cm between them, an omni-directional sound sensor ( $SI$ ), and a loud-speaker (see Figure 2 right). The motion of the robot implemented by the two wheel actuators ( $M_1$  and  $M_2$ ) is simulated by the Differential Drive Kinematics equations, as presented in [3]. Light and proximity sensor values are simulated through a sampling technique [7]. The robot floor sensors output the following values: 0 if the robot is positioned over white floor;  $\frac{1}{3}$  if the robot is positioned over light grey floor;  $\frac{2}{3}$  if the robot is positioned over dark grey floor; 1 if the robot is positioned over black floor. The speaker is simulated as producing a binary output (on/off); the sound sensor has no directionality and intensity features. 10% uniform noise was added to the light and proximity sensor readings, the motor outputs and the position of the robot. We also added noise of 5% on the reading of the two floor sensors, by randomly flipping between the 4 aforementioned values, in order to deal with wrong readings produced by the real robot while it is moving.

**The controller and the evolutionary algorithm** Fully connected, thirteen neuron CTRNNs are used. All neurons are governed by the following state equation:

$$\frac{dy_i}{dt} = \frac{1}{\tau_i} \left( -y_i + \sum_{j=1}^{13} \omega_{ji} \sigma(y_j + \beta_j) + gI_i \right), \quad \sigma(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

where, using terms derived from an analogy with real neurons,  $y_i$  represents the cell potential,  $\tau_i$  the decay constant,  $\beta_j$  the bias term,  $\sigma(y_j + \beta_j)$  the firing rate,  $\omega_{ji}$  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  $i$ . Nine neurons receive input from the robot sensors: i.e., neuron  $N_1$  takes input from  $\frac{L_1+L_2}{2}$ ,  $N_2$  from  $\frac{L_3+L_4}{2}$ ,  $N_3$  from  $F_1$ ,  $N_4$  from  $F_2$ ,  $N_5$  from  $\frac{P_1+P_2+P_3+P_4}{4}$ ,  $N_6$  from  $\frac{P_5+P_6+P_7+P_8}{4}$ ,  $N_7$  from  $\frac{P_9+P_{10}+P_{11}+P_{12}}{4}$ ,  $N_8$  from  $\frac{P_{13}+P_{14}+P_{15}}{3}$ , and  $N_9$  from  $SI$ . The neurons  $N_1$  to  $N_8$  receive as input a real value in the range  $[0,1]$ , while the neuron  $N_9$  receives a binary input (i.e., 1 if a tone is emitted, 0 otherwise). The other neurons do not receive any input from the robot's sensors. The cell potential ( $y_i$ ) of the

11<sup>th</sup> neuron, mapped into [0,1] by a sigmoid function ( $\sigma$ ) is used by the robot to control the sound signalling system (i.e., the robot emits a sound if  $y_{11} \geq 0.5$ ). The cell potential ( $y_i$ ) of the 12<sup>th</sup> and the 13<sup>th</sup> neuron, mapped into [0,1] by a sigmoid function ( $\sigma$ ) and then linearly scaled into [-4.0,4.0], set the robot motors output. The strength of synaptic connections  $\omega_{ji}$ , the decay constant  $\tau_i$ , the bias term  $\beta_j$ , 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 (GA) is employed to set the parameters of the networks [4]. The population contains 100 genotypes. Each genotype is a vector comprising 196 real values (169 connections, 13 decay constants, 13 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]. 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. 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 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  $\tau_i$  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 the evolution, each genotype is coded into a robot controller, and is evaluated for 15 trials, 12 trials in *Env A*, and 3 trials in *Env B*, sequentially. The latter choice stems from the fact that a bigger proportion of *Env A* seems to favor the evolution of a robust discrimination mechanism [15]. Each trial  $e$  differs from the others in the initialisation of the random number generator, which influences the robot’s 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 110 s (1100 simulation cycles). In each trial, the robots are rewarded by an evaluation function  $f_e$  which is given by the following formula:

$$f_e = \frac{\sum_{i=1}^2 f_i}{2}, \quad f_i = \frac{C_p + C_a + C_s}{n_c}, \quad i = 1, 2, \quad \text{where :}$$

- 1)  $C_p$  rewards phototaxis, that is movement to the target area:

$$C_p = \frac{d_i - d_c}{d_i}$$

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*,  $C_p$  is set to 1 if the robot terminates the trial in the target area. In *Env B*,  $C_p$  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)  $C_a$  rewards movements away from the light (anti-phototaxis):

$$C_a = \begin{cases} \frac{d_c}{d_{max}} & \text{if trial in } Env B, C_p = 1 \text{ and } SI = 1; \\ 0 & \text{otherwise} \end{cases}$$

where  $d_{max} = 110$  cm is the maximum distance a robot can reach to the light.

3)  $C_s$  rewards agents that (i) never signal in *Env A*; (ii) signal in *Env B*. In *Env B* if one robot signals, both are rewarded.

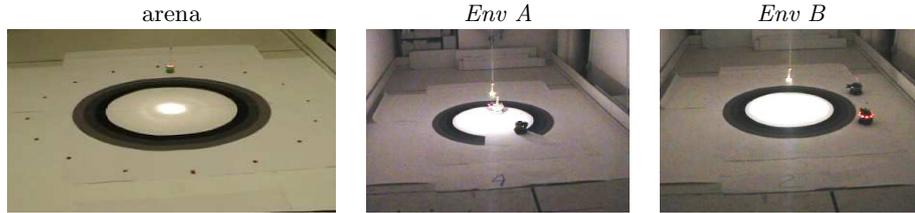
$$C_s = \begin{cases} 1 & \text{if proper signalling} \\ 0 & \text{otherwise} \end{cases}$$

4)  $n_c$  is the number of crashes in a trial (the number of times the robots get closer than 2.5 cm). If  $n_c > 3$ , the trial is considered finished.

An important feature of this evaluation function is that it simply rewards agents that make proper use of their sound signalling system, without directly interfering with the nature of the discrimination strategies. Last but not least, we do not explicitly code a reward for the use of communication, nevertheless, as will be shown in Section 3, there is an indication that implementation details of our experimental setup lead to a selective pressure in favour of the evolution of communication.

### 3 Results

Ten evolutionary simulations, each using a different random initialisation were run for 10000 generations. Three of those runs were successful, therefore evolved the required behaviour: in *Env A*, both robots reach the target area without emitting any sound signal and in *Env B* (a) they reach the circular band and start circling around the light looking for a *way in* zone, (b) after performing approximately one tour around the light, they signal the absence of the *way in* zone by emitting a tone (c) as a reaction to the sound signal (emitted or perceived), they leave the band reaching a desired distance from the light. The rest of the runs did not succeed to evolve the discrimination mechanism and robots were never signalling in *Env B*. Therefore we succeeded in our first goal which was to evolve the desired behaviour and obtain a controller capable of solving the complex task described in Section 2. Given the fact that we use a single integrated (i.e. non-modularised) network, we have the feeling that the obstacle avoidance behaviour makes it significantly harder for evolution to solve the discrimination task. Robot-robot interactions might require behaviour phylogenetically antecedent to other more complex responses. The evolution of such behaviour might be based on some mechanism that does not facilitate the subsequent evolution of other neural structures required by the robot to successfully perform other parts of the task (i.e. integration over time). This hypothesis though demands further investigation in order to be confirmed, but might explain why only 30% of the runs were successful.

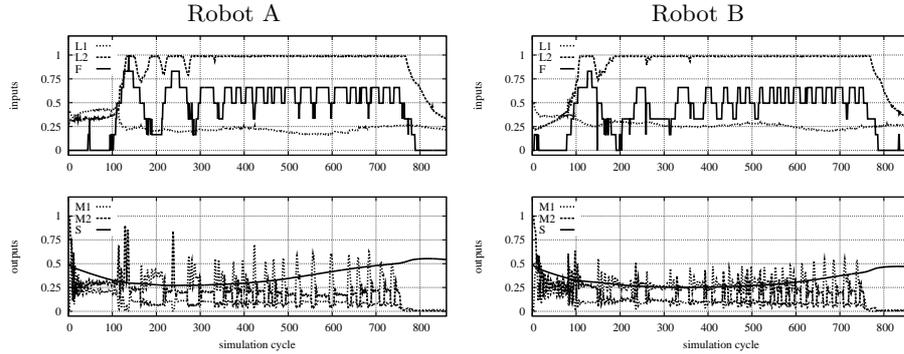


**Fig. 3.** The experimental setup. Left: a picture of the arena, with the points around the band showing the positions in which the robots were randomly initialised. Middle: a snapshot of a trial in which two robots find the *way in zone* in *Env A*. Right: a snapshot of a trial in *Env B*. The robot with the lighter turret colour is the one that has signalled the absence of a *way in zone*, while both robots have left the band performing anti-phototaxis.

**Experiments with real *s-bots*** As already pointed out in the introduction, the aim of this paper is to test and evaluate the evolved controllers on the real robots. A major issue is the choice of the genotype which will control the robots. Since all genotypes of the latest generations of the three successful runs achieve the maximum fitness score, we had to come up with some more specific criteria in order to choose our controller. These criteria are plainly good sensory-motor coordination and effective sound signalling behaviour. These properties are quite obvious since they guarantee a better overall performance and better visualisation of the solution, but also seem to be a prerequisite in order to avoid misreadings from the robot’s ground sensors. Therefore, after testing various controllers on the real robots, we made our choice according to these criteria and only. We performed 40 trials, 20 for each environment, in the arena shown in Figure 3, left<sup>2</sup>. The robots were initialised at a distance of 85 cm from the light with a random orientation. In *Env A* we randomly varied the position of the *way in zone* but we fixed its width to 45°, which is the smallest value encountered through evolution and the most difficult case for a possible misinterpretation of an *Env A* for an *Env B*. The results for the 40 trials performed are 100% successful: that is the robots never make any type of error, like crossing the black edge of the band, crashing against each other, erroneously signalling in *Env A* or failing to signal in *Env B*. This is an extremely successful result since it is achieved on the real hardware.

**Behavioural Analysis** At this point we will give a description of the behaviour exhibited by the robots trying to get an insight on how they come to solve the task described in Section 2. The mechanism employed by the robots is displayed in Figure 4 and is similar to what described in [15]. The robots encounter the band at around 100 timesteps. At that point their sound output starts to increase with a rather constant slope. In *Env A* it starts decreasing before reaching the threshold of 0.5 once the *way in zone* is encountered. On the contrary, in *Env B*, shortly after 750 timesteps it goes beyond the threshold of 0.5 for robot *A* and that robot signals, while robot *B* does not. After that, light and floor readings of both robots decrease since the robots leave the band performing anti-phototaxis.

<sup>2</sup> The movies that correspond to the 40 trials can be found in <http://iridia.ulb.ac.be/~campatzis>



**Fig. 4.** The behavioural analysis of the two real robots in the same trial in *Env B*. Left: the inputs and outputs of robot *A* for 1000 timesteps. Right: same for robot *B*. See text for details

We can clearly see the effect of communication for robot *B*, since it leaves the band without having signalled. Concerning the anti-phototaxis behaviour, we can deduce from Figure 4 that the mechanism for both robots is simply putting both motor outputs at a value near 0, which corresponds to moving almost straight backwards at full speed<sup>3</sup>.

To further analyse our results we need to introduce a measure of the difference between the entrance position of the robot in the circular band and the position at which the robot starts to signal. This measure, called offset  $\Delta$ , is computed as follows:

$$\Delta = |\alpha(t_e, t_s)| - 2\pi, \quad \alpha(t_1, t_2) = \sum_{t=t_1}^{t_2-1} \widehat{\mathbf{AOB}}, \quad \mathbf{A} = \mathbf{X}_t, \mathbf{B} = \mathbf{X}_{t+1} \quad (2)$$

where  $\mathbf{O}$  corresponds to the position of the light, and  $\alpha$  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.  $\alpha$  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. Offset  $\Delta$  takes value  $0^\circ$  if the robot signals exactly after covering a complete loop of the circular band. Negative values of the offset  $\Delta$  suggest that the robot signals before having performed a complete loop, while positive values correspond to the situation in which the robot emits a tone after having performed a loop around the light.

Looking at our results, we observed that in *Env B* it was always the same robot (*A*) that was signalling first, regardless of the robots' initialisation. Since the integration of the light and floor sensors perception is used to make the discrimination, and given the differences in the readings of the two robots (see Figure 4), we can attribute the fact that robot *A* always signals earlier than *B* to mechanical or sensor implementation differences. On average, as also shown in

<sup>3</sup> In all experiments the robots are moving backwards—in a opposite direction of their gripper

**Table 1.** Quantitative data about the performance of the system. We compare the 20 trials in *Env B* with the real robots with 500 evaluations in simulation, for the same random initialisation methodology. We measure the average offset  $\Delta$  of the first signalling robot in each trial—since the system uses communication—and the reaction time of the signalling and listening robot (respectively  $t_{react,s}(s)$  and  $t_{react,l}(s)$ ). We compare them to the corresponding values obtained in simulation.

	Offset $\Delta$		$t_{react,s}(s)$		$t_{react,l}(s)$	
	avg	sd	avg	sd	avg	sd
Real Robots	-30.6	11.75	8.78	0.75	7.69	0.57
Simulation	+31.6	16.05	9.25	1.1	9	1.66

Table 1, we see that robot *A* signals *before* completing a loop, but given the fact that our *way in zone* is fixed to  $45^\circ$ , its behaviour is still safe and does not lead to possible misinterpretations of *Env A* (thus signalling in those environments). In a separate test we let robot *B* (which never signalled first) perform the task alone in order to get an estimate of its offset<sup>4</sup>. We observed that it was always signalling *after* completing a loop. However, thanks to the use of communication between the two agents, both robots leave the band once the first robot signals. The result is that the system is more effective than in the case where each robot performs the task by itself, since in that case the second robot would signal much after the first and thus would conclude the task with a considerable temporal delay. As a comparison, we performed a test to see when the robots signal in simulation. We re-evaluated the same genotype used in our real robot experiments 500 times for *Env B* in exactly the same conditions as in reality (same initialisation, orientation). As shown in Table 1, the robots signal *after* completing the loop, and the behaviour of the real robots is slightly different from the one of the simulated ones. The conclusion is that the two real robots behave in a different way as a result of their sensorial differences. In our case the behaviour was not disrupted by those differences with respect to simulation, but in case our system presented noise far beyond the level encountered during evolution, we might have errors like wrong discrimination or very late signalling.

We have also measured the reaction time of the system with respect to the discrimination signal in *Env B*. As shown in Table 1, we observe that on average the robot that hears the signal is reacting faster than the one that emits sound, that is leaves the band and reaches a certain distance (85 cm) faster. Obviously, the two controllers are in a different state once the signal is emitted, and that could produce the difference in the reaction time, but also mechanical differences of each robot could influence this behaviour. The difference though in simulation is significantly smaller therefore it is very hard to infer any conclusions.

**Evolution of communication** We are also interested in investigating the evolutionary significance and the effect of communication. Specifically, we want to study the evolution of communication in a system where there is no explicitly coded fitness advantage from it. To do so, we tested in simulation the behaviour of two genotypes from the same run, that both score the maximum. As we will prove, one (*G1*, of an earlier generation) is relying on robots that solve the task using the information they themselves collect from their environment, while the other (*G2*, of a later generation) also relies on communication. To prove that

<sup>4</sup> Data not shown.

**Table 2.** Results of 500 re-evaluations in *Env B* for the genotype with (*G2*) and without (*G1*) evolved reaction to the communication signal. We compare the average and standard deviation of the task completion time of both robots—one initialised close to the light and one further away.

	Communication OFF				Communication ON			
	$t_{react\ s-bot\ near}$ (s)		$t_{react\ s-bot\ far}$ (s)		$t_{react\ s-bot\ near}$ (s)		$t_{react\ s-bot\ far}$ (s)	
	avg	sd	avg	sd	avg	sd	avg	sd
<i>G1</i>	98.99	11.045	113.3	8.23	100.51	10.13	114.74	7.36
<i>G2</i>	99.33	8.03	108.09	6.18	99.5	7.38	99.15	8.53

communication can indeed influence the performance of the system, we came up with the following setup: we initialised the two robots at two different positions: one closer (75cm) to the light and one further away (95 cm) and performed two tests with each genotype. In one test, the communication channel is “open”, so the robots can hear the signal of the other robot, and react to it (*G2*) or ignore it (*G1*). In the other test we deliberately “close” the channel by not allowing each robot to listen to the signals of the other one. Thus, to what concerns the genotype not relying on communication, there should be no difference while the genotype exploiting communication should show different behaviours. By introducing the difference in the initial position, we make sure that on average one robot will signal earlier than the other. Thus we will prove that (a) the first genotype indeed displays no communication properties while the second does, and therefore communication has evolved and (b) communication can have a performance effect on the system, in this case making it more efficient. We perform the above re-evaluation for 500 times in *Env B*, for each setup (communication channel on/off) and for both genotypes. Indeed, looking at Table 2 we see that the time it takes for the robot initialized far away to reach the desired distance (set to 85 cm) from the light is significantly reduced when communication is used, while closing the channel in the case of the first genotype has no effect on this measure. A crucial question is why did communication evolve since there was no explicitly coded fitness gain for it, and under which circumstances. We observe that communication evolved in our system *after* the two robots evolved the individual discrimination mechanism. Due to some random mutation at that point, the robot that signals later started to react to the signal of the other robot, and that solution spread in the population, because the system cannot do worse when communication is active. To prove that we ran a statistical analysis. *G2* was evaluated for 15000 times in simulation, 12000 in *Env A* and 3000 times in *Env B*, keeping the proportion encountered during evolution, in two setups: one setup with the communication channel open, and another with it closed. We performed a pair-wise Wilcoxon test, which confirmed with a 99% confidence that for this particular individual, the communicative setup performed constantly better than the non-communicative one. Since the evaluation function as presented in Section 2 does not explicitly reward communicating agents, the results of the test suggest that it is through the implementation details of the experimental setup (initialisation of the robots, noise in the system, the choice when to terminate a trial etc.) that the use of communication ends up to be advantageous during the evolutionary process.

## 4 Conclusions

The main challenge of the work presented was to evaluate a solution obtained by artificial evolution on real robots. Our results show that non-reactive mechanisms for memory and learning, evolved in a simulated world, can deal with reality. The “coarse” nature of our simulator—no dynamics, friction or momentum modelled—allowed us to find effective solutions to a complex task, without loss of performance and gaining in computational time. Owing to these promising results, we can further exploit our approach in order to (i) investigate the mechanisms involved in complex non-reactive tasks and (ii) pursue our studies on the evolution of communication in embodied and situated agents.

## References

1. R.A. Brooks. Artificial life and real robots. In F.J. Varela and P. Bourguine, editors, *Towards a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life*, pages 3–10. MIT Press, Cambridge, MA, 1992.
2. G. De Croon, S. Nolfi, and E.O. Postma. *Toward pro-active embodied agents: On the importance of neural mechanisms suitable to process information in time*. Complex Engineering Systems, Perseus Books Groups Press, 2004. In Press.
3. G. Dudek and M. Jenkin. *Computational Principles of Mobile Robotics*. Cambridge University Press, Cambridge, UK, 2000.
4. D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA, 1989.
5. N. Jakobi. Evolutionary robotics and the radical envelope of noise hypothesis. *Adaptive Behavior*, 6:325–368, 1997.
6. M.J. Matarić and D. Cliff. Challenges in evolving controllers for physical robots. *Robotics and Autonomous Systems*, 19(1):67–83, 1996.
7. O. Miglino, H.H. Lund, and S. Nolfi. Evolving mobile robots in simulated and real environments. *Artificial Life*, 2(4):417–434, 1995.
8. F. Mondada, G. C. Pettinaro, A. Guignard, I. V. Kwee, D. Floreano, J.-L. Deneubourg, S. Nolfi, L. M. Gambardella, and M. Dorigo. SWARM-BOT: A new distributed robotic concept. *Autonomous Robots*, 17(2–3):193–221, 2004.
9. S. Nolfi and D. Floreano. *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MIT Press, Cambridge, MA, 2000.
10. E.A. Di Paolo. Spike-timing dependent plasticity for evolved robots. *Adaptive Behavior*, 10(3–4):243–263, 2003.
11. V. Trianni, S. Nolfi, and M. Dorigo. Cooperative hole avoidance in a *swarm-bot*. *Robotics and Autonomous Systems*, 2005. In press.
12. E. Tuci. *An exploration on the evolution of learning behaviour using robot based models*. PhD thesis, University of Sussex, Brighton, UK, 2004.
13. E. Tuci and M. Quinn. Behavioural plasticity in autonomous agents: a comparison between two types of controller. In Günther Raidl et al., editor, *Proceedings of the 2<sup>nd</sup> European Workshop on Evolutionary Robotics (EvoROB2003)*, pages 661–672. Springer-Verlag, 2003.
14. E. Tuci, M. Quinn, and I. Harvey. An evolutionary ecological approach to the study of learning behaviour using robot-based model. *Adaptive Behavior*, 10(3–4):201–221, 2003.
15. E. Tuci, V. Trianni, and M. Dorigo. ‘Feeling’ the flow of time through sensory/motor coordination. *Connection Science*, 16:1–24, 2004.