Evolution of Signalling in a Group of Robots Controlled by Dynamic Neural Networks

Christos Ampatzis[1], Elio Tuci[1], Vito Trianni[2], and Marco Dorigo[1]

[1] IRIDIA, CoDE, Université Libre de Bruxelles,
{campatzi,etuci,mdorigo}@ulb.ac.be
http://iridia.ulb.ac.be/
[2] ISTC-CNR, Rome, Italy
vito.trianni@istc.cnr.it

Abstract. Communication is a point of central importance in swarms of robots. This paper describes a set of simulations in which artificial evolution is used as a means to engineer robot neuro-controllers capable of guiding groups of robots in a categorisation task by producing appropriate actions. In spite of the absence of explicit selective pressure (coded into the fitness function) which favours signalling over non-signalling groups, communicative behaviour emerges. Post-evaluation analyses illustrate the adaptive function of the evolved signals and show that they are tightly linked to the behavioural repertoire of the agents. Finally, our approach for developping controllers is validated by successfully porting one evolved controller on real robots.

1 Introduction

Recently, there has been a growing interest in multi-robot systems since, with respect to a single robot system, they provide increased robustness by taking advantage of inherent parallelism and redundancy. Moreover, the versatility of a multi-robot system can provide the heterogeneity of structures and functions required to undertake different missions in unknown environmental conditions. Among the possible theoretical perspectives which currently guide the design of multi-robot systems, the swarm robotics approach is characterised by its emphasis on aspects such as decentralisation of the control, limited communication abilities among robots, use of local information, emergence of global behaviour and robustness [1].

Given a multi-robot system with such properties, a global distributed knowledge of, for example, the status of the environment, can be achieved by exploiting the local knowledge of each single robot and by propagating the latter through various forms of communication. For this reason, research in swarm robotics dedicates particular attention to the study of how local information can be efficiently communicated among the robots, so to improve the adaptiveness of the group (see [2]). In this paper, we describe a simulation work in which we provide a group of two robots with a sound signalling system (i.e., "ears" and "mouth") and we investigate the conditions which favour the emergence

of a communication protocol. In particular, our work studies the evolution of signalling in a group of autonomous robots within the context of *decision making* and *action selection*, where robots have to make decisions by categorising their environment and perform different actions. The categorisation of the environment results from how the robots' sensory inputs unfold in time (see [3,4] for similar examples). The decision making is performed at the individual level, and a collective action should be the observed response to the individual decision.

In order to perform our study, we make use of the research method referred to as Evolutionary Robotics (ER, see [5]). Roughly speaking, ER is a methodological tool to automate the design of robots' controllers. Based on artificial evolution, ER finds sets of parameters for artificial neural networks (ANN's) that guide the robots to the accomplishment of their objective. ER can be employed to look at the effects that the physical interactions among embodied agents and their world have on the evolution of individual behaviour and social skills (see [6]). ER also permits the co-evolution of communicative and non-communicative behaviour, since it lets different characteristics co-adapt, only requiring an overall evaluation of the group (see [7]). Note that one of the main features of this work is that we do not explicitly reward the group for displaying signalling behaviour. That is, the adaptive pressure coded into the fitness function does not explicitly favour signalling over non-signalling groups. Therefore, if the evolved robot controllers display any kind of signalling behaviour, the adaptive significance of this feature has to be investigated. The reason to entirely leave the development of communicative behaviour to artificial evolution resides in the fact that in this way the co-adaptation of all mechanisms can produce more effective ways to categorise sensory-motor information. Evolution can produce solutions better adapted to the problem than hand-coded signalling behaviour (see [2]).

Our aim is to examine the evolution of communication in a group of homogeneous robots, in close relation to the mechanisms that govern the robots' behaviour with respect to the task. We will show that communication is beneficial for the group and that its adaptive function is tightly connected to *action selection* and *decision making*. Finally, we will download the evolved controllers on real robots, which is the only way to prove the validity of the chosen design methodology (i.e., artificial evolution). Even though multiple works treat the issue of porting a non-reactive controller to reality, the literature lacks works addressing tasks where the integration over time of sensory input is required. In these cases, the decision making relies on how the inputs unfold in time and possible errors will accumulate through time and could severely disrupt the performance. In what follows, we describe the task (Section 2), the simulation model (Section 3), the controller and the evolutionary algorithm (Section 4), and the fitness function employed (Section 5). Results in simulation are presented in Section 6, results on real hardware are discussed in Section 7 and conclusions are drawn in Section 8.

2 Description of the task

At the start of each trial, two simulated robots are placed in a circular arena with a radius of 120 cm (see Fig. 1), at the centre of which a light bulb is always turned on. The robots are positioned randomly at a distance between 75 and 95 cm from the light, with a random orientation between $-120°$ and $+120°$ with respect to it. 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 covering an area between 40 and 60 cm from it. 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. Robots are not allowed to cross the black edge of the band close to the light. 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. 1a). In the other type, referred to as *Env B*, the band completely surrounds the light (see Fig. 1b). The *way in* zone represents the path along which the robots are allowed to safely reach the *target area* in *Env A*—an area of 25 cm around the light. 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 and reach a certain distance from the light source. Robots have 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 "recognising" the absence of the *way in* zone and leave the band by performing antiphototaxis. Artificial evolution is used to design controllers capable of providing the robot with the mechanisms required to solve the task.

Each robot is required to use a temporal cue in order to discriminate between *Env A* and *Env B*, as in [4]. This discrimination is based on the persistence of the perception of a particular sensorial state (the floor, the light or both) for the

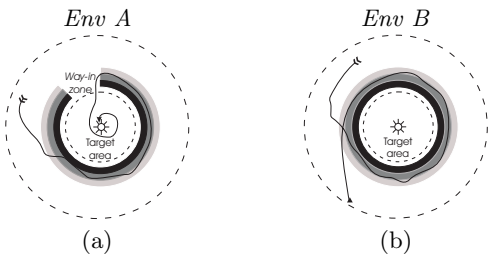

Fig. 1. The task. (a) *Env A* is characterised by the *way in* zone. The *target area* is indicated by the dashed circle. (b) In *Env B* the target area cannot be reached. The continuous arrows are an example of a good navigational strategy for one robot.

(a) (b)

Fig. 2. (a) A picture of an *s-bot*. (b) Plan of the simulated robot, 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 (see text for details). The wheel motors are indicated by M_1 and M_2. A sound signalling system (loud speaker) is referred to as S.

amount of time that, given the trajectory and speed of the robot, corresponds to the time required to make a loop around the light. The integration over time of the robots' sensorial inputs is used to trigger antiphototaxis in *Env B*.

Communication is not required to solve the task described above. However, robots are provided with a sound signalling system that can be used for communication. Given that we provide the agents with "mouth" and "ears", whenever a robot produces a signal ("talker"), this signal is "heard" by itself and the other agent. The fitness function we use does not explicitly reward the use of signalling. We investigate whether or not the latter evolves and in case it does, what its adaptive function is. Finally, we use a homogeneous group of robots, that is the same neural controller is cloned on both robots.

3 The simulation model

The controllers are evolved in a simulation environment which models some of the hardware characteristics of the *s-bots* (see Fig. 2a). The *s-bots* are small wheeled cylindrical robots, 5.8 cm of radius, equipped with a variety of sensors, and whose mobility is ensured by a differential drive system [8]. In this work, we make use of four ambient light sensors, placed at $-112.5°$ (L_1), $-67.5°$ (L_2), $67.5°$ (L_3), and $112.5°$ (L_4) with respect to its heading, fifteen infra-red proximity sensors placed around its turret (P_1 to P_{15}), two floor sensors F_1 and F_2 positioned facing down 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 Fig. 2b). 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 [9]. Light and proximity sensor values are simulated through a sampling technique. 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

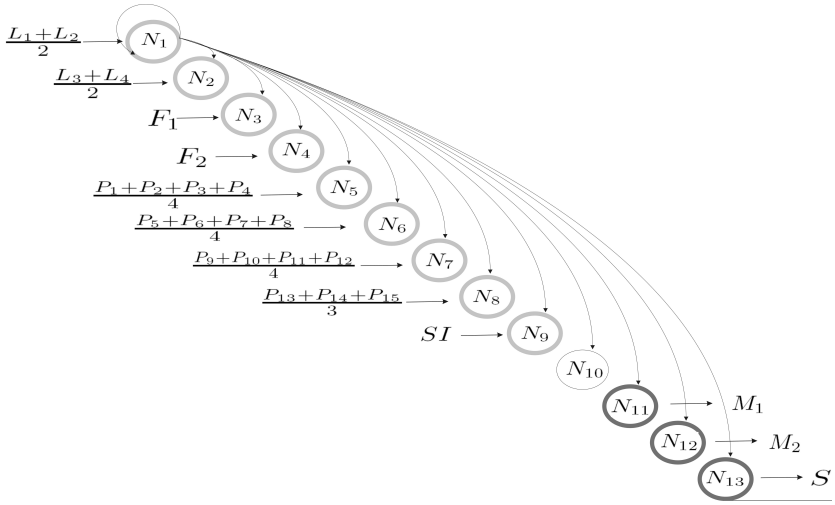

Fig. 3. The fully connected CTRNN architecture. Only the efferent connections for N_1 are drawn and all neurons behave in the same way. Neurons are represented as circles. Circles with the light grey outline represent the input neurons, while circles with the heavy grey outline represent the output neurons. We show for all input neurons the combination of sensors that serve as inputs, and for all output neurons the corresponding actuator. N_{10} is not connected to any sensor or actuator.

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. During evolution, 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.

4 The controller and the evolutionary algorithm

Given that the task we want to study requires the use of time-dependent structures, we use fully connected, thirteen neuron Continuous Time Recurrent Neural Networks (CTRNN's see [10])—see Fig. 3 for a depiction of the network. All neurons are governed by the following state equation:

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

where, using terms derived from an analogy with real neurons, τ_i is the decay constant, y_i represents the cell potential, ω_{ji} the strength of the synaptic connection from neuron j to neuron i, $\sigma(y_j + \beta_j)$ the firing rate, β_j the bias term, g the gain and I_i the intensity of the sensory perturbation on sensory

neuron i. The connections of all neurons to sensors and actuators is shown in Fig. 3. Neurons N_1 to N_8 receive as input a real value in the range [0,1]. Neuron N_1 takes as input $\frac{L_1+L_2}{2}$, $N_2 \leftarrow \frac{L_3+L_4}{2}$, $N_3 \leftarrow F_1$, $N_4 \leftarrow F_2$, $N_5 \leftarrow \frac{P_1+P_2+P_3+P_4}{4}$, $N_6 \leftarrow \frac{P_5+P_6+P_7+P_8}{4}$, $N_7 \leftarrow \frac{P_9+P_{10}+P_{11}+P_{12}}{4}$ and $N_8 \leftarrow \frac{P_{13}+P_{14}+P_{15}}{3}$.

Neuron N_9 receives a binary input (i.e., 1 if a tone is emitted, 0 otherwise) from the microphones of both robots (SI), while N_{10} does not receive input from any sensor. The cell potentials (y_i) of N_{11} and N_{12}, mapped into [0,1] by a sigmoid function (σ) and then linearly scaled into [-4.0,4.0], set the robot motors output. The cell potential of N_{13}, mapped into [0,1] by a sigmoid function (σ) is used by the robot to control the sound signalling system (the robot emits a sound if $y_{13} \geq 0.5$). The strength of the 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 [11]. The population contains 100 genotypes. Generations after the first are produced by a combination of selection with elitism, recombination and mutation. More details on the evolutionary algorithm employed and on the genotypes' component values can be found in [12].

5 The fitness function

During evolution, each genotype is coded into a robot controller, and is evaluated for 10 trials, 5 in each environment. The sequence order of environments within the ten trials has no bearing on the overall performance of the group since each robot controller is reset at the beginning of each trial. Each trial 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 (between 45° to 81°), and the noise added to motors and sensors. Within a trial, the robot life-span is 100 s (1000 simulation cycles). The final fitness attributed to each genotype is the average fitness score of the 10 trials. In each trial, the fitness function E is given by the following formula:

$$E = \frac{E_1 + E_2}{2 * (n_c + 1)}$$

where n_c is the number of (virtual) collisions in a trial, that is the number of times the robots get closer than 2.5 cm to each other (if $n_c > 3$, the trial is terminated) and E_1 and E_2 are the fitness scores of each robot i, calculated as follows:

$$E_i = \begin{cases} \frac{d_i - d_f}{d_i} & \text{if during } \textit{phototaxis}\text{, if } \textit{cross} \text{ or trial in } \textit{Env A} \\ 1 + \frac{d_f - 40}{d_{max} - 40} & \text{if band is reached in } \textit{Env B} \end{cases}$$

By d_i we refer to the distance from the light at which the robot was initialised and d_f is the distance from the light at which the robot is at the end of the trial

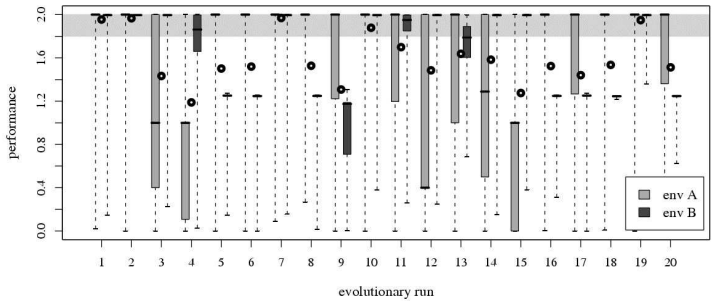

Fig. 4. Box-and-whisker plot visualising the post-evaluated fitness of groups a1-a20 in both environments. The box comprises observations ranging from the first to the third quartile. The median is indicated by a horizontal bar. When the observations are too close the box degenerates to the median. The whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range. The gray area denotes the area into which the average fitness value for both environments (black circles) must be, in order for the group to be called successful.

and $d_{max} = 120$ cm is the maximum distance from the light a robot can reach. By *phototaxis* we mean the phase where robots in either environment have not yet touched the band in shades of grey. By *cross* we mean crossing the black edge of the band in either environment. In case $robot_i$ ends up in the *target area* in *Env A*, we set $E_i = 2$. From the above equation we can see that this is also the maximum value E_i can get for a robot in *Env B*, and this corresponds to the robot ending up at 120 cm from the light ($d_f = 120$). So if both robots are successful, the trial gets the maximum score of 2. An important feature of this fitness function is that it rewards agents that develop successful discrimination strategies and end up doing the correct action in each environment, regardless of any use of sound signalling. That is, a genotype that controls a group that solves the task without any signalling/communication gets the same fitness as one that makes use of communication.

6 Results

Twenty evolutionary simulations, each using a different random initialisation, were run for 12000 generations. It is important to note that the fitness of the best evolved controllers during evolution may have been an overestimation of their ability to guide the robots in the task. In general, the best fitness scores take advantage of favourable conditions, which are determined by the existence of between-generation variation in starting position and orientation and other simulation parameters. In order to have a better estimate of the behavioural capabilities of the evolved controllers, we post-evaluate, for each run, groups

controlled by the best genotype of the last generation. Groups controlled by neural networks built by those genotypes will be from now on referred to as a1-a20. The entire set of post-evaluations (500 trials in *Env A* and 500 trials in *Env B*) should establish whether (i) a group of robots can solve the task (ii) a sound signalling mechanism has been evolved and what its functionality is. The results of the post-evaluation phase are shown in Fig. 4. We define as successful a genotype that after this phase has an average fitness value above 1.8. This roughly corresponds to both robots reaching the *target area* in *Env A* and leaving the band performing anti-phototaxis in *Env B*. The results suggest that five of the groups produced satisfying solutions to the task (a1, a2, a7, a10, a19). Table 1 shows that four of the successful groups (a1, a2, a7, a19) make large use of signalling in *Env B*, while in *Env A* signalling is negligible—see column 6 and 8, which refer to the average percentage of time either robot emits a signal during a trial. Among the successful groups only a10 did not use signalling. To unveil the relationship between the emission of sound signals and the completion of the task we perform a behavioural analysis of the successful evolutionary runs. Thus, we evaluate the five successful groups in a different setup in which the robots are not able to perceive any sound from the environment: their sound input is set to 0 all the time. We refer to this condition as the *deaf* setup. The results of this analysis are shown in Table 1, together with the results for the *normal* setup—i.e., without applying any disruption. The first observation that we make is that, for group a10, the average fitness in the *deaf* setup (see Table 1 columns 10, 12) is exactly the same as the one in the *normal* setup (see Table 1 columns 2, 4). For the other groups, the average fitness for *Env B* drops considerably in the *deaf* setup, while for *Env A* it remains approximately the same. This suggests that group a10 does not rely on the activation of the sound input in order to solve the task in *Env B* while the other groups do. Furthermore, for

Table 1. Further results of post-evaluation tests with *normal* and *deaf* setup for the five successful groups. For the *normal* setup, the table shows: (i) the average and standard deviation of the fitness over 500 trials in *Env A* (see columns 2, and 3) and in *Env B* (see columns 4, and 5); (ii) the average and standard deviation of the percentage of time-steps the sound was on by either robot over 500 trials in *Env A* (see columns 6, and 7) and in *Env B* (see columns 8, and 9). For the *deaf* setup the table shows the average and standard deviation of the fitness over 500 trials in *Env A* (see columns 10, and 11) and in *Env B* (see columns 12, and 13).

group	*normal*								*deaf*			
	fitness				*signalling (%)*				*fitness*			
	Env A		*Env B*		*Env A*		*Env B*		*Env A*		*Env B*	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
a1	1.927	0.310	1.982	0.134	0.03	0.57	21.62	3.82	1.937	0.292	1.0526	0.239
a2	1.937	0.277	1.995	0.002	0.77	4.49	18.48	1.17	1.969	0.156	1.256	0.094
a7	1.988	0.113	1.950	0.198	0	0	17.19	2.54	1.988	0.113	1.266	0.250
a10	1.789	0.467	1.968	0.183	0	0	0	0	1.789	0.467	1.968	0.183
a19	1.914	0.236	1.984	0.059	0.08	0.72	13.88	1.03	1.923	0.214	1.137	0.016

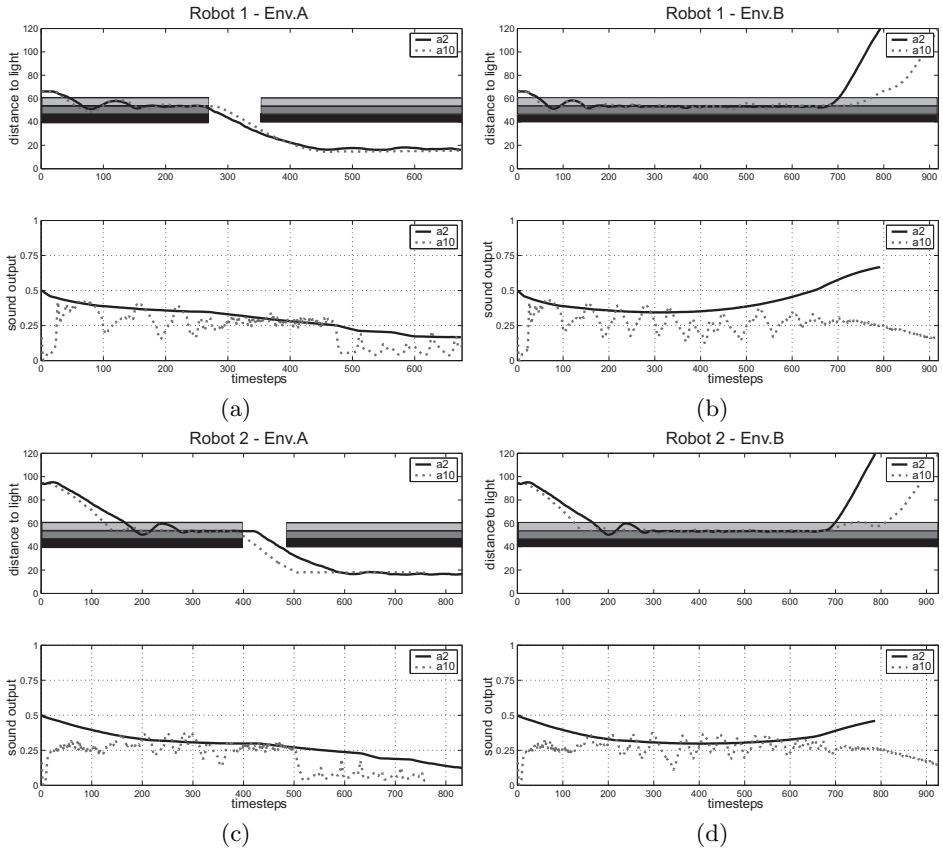

Fig. 5. The graphs show some features of the behaviour of robots of groups a2 (continuous lines) and a10 (dashed lines), during a successful trial in *Env A* and in *Env B*. Top graphs a, b, c, and d show the distance to the light in cm. Bottom graphs a, b, c, and d show the firing rate of neuron N_{13} (i.e., the sound output) of each robot controller.

the latter groups, the fitness value in the *deaf* setup in *Env B* corresponds to both robots not performing antiphototaxis; that is, the robots stay on the band and keep circling around the light. In order to understand the function of these signals, we looked more carefully at the behaviour of groups a2 and a10 during a successful trial in each environment. In particular, our analysis focuses on the relationship between the robot-light distances and the firing rate of neuron N_{13} of each controller of a group, since this neuron triggers the emission of sound. Fig. 5a, b, c, d (top) show the distances of each robot to the light at every timestep. The areas in shades of grey in these graphs represent the circular band. Fig. 5a, b, c, d (bottom) show the firing rate of N_{13} (i.e., the sound output) of both robots of a group. In all graphs, continuous lines refer to robots of group a2, dashed lines refer to robots of group a10.

As shown in Fig. 5a, b, c, d (top), the behaviour of the robots can be divided in three phases. In the first two phases the robots of both groups (a2 and a10) behave in the same way in both environments. The robot-light distance initially decreases up to the point where the robots touch the band (phototaxis phase) and then stays quite constant as the robots circle around the band trying to find the *way in* zone (integration over time phase). In the third phase the groups behave differently according to the characteristics of the environment. For both groups, in *Env A* the robot-light distance decreases further as the robots end up in the *target area*, while in *Env B* it increases and reaches the maximum distance as the robots leave the band (antiphototaxis phase). Concerning the firing rate of neuron N_{13}, the two groups differentiate in both environments (see Fig. 5a, b, c, d (bottom)). In *Env B*, the firing rate for both robots of group a10, never goes beyond the threshold of 0.5 (see Fig. 5b, d (bottom) dashed lines). In the case of a2 though, the firing rate of N_{13} of Robot 1 is rising until it passes over the threshold of 0.5 just before the robot starts performing antiphototaxis (see Fig. 5b, d, continuous lines). This behaviour of N_{13} for a2 reflects the integration over time process, which leads to passing over the threshold of 0.5 in *Env B* (decision making), while it is interrupted in *Env A*, when the *way in* zone is found (see Fig. 5a, c, continuous lines). Differently, for a10 this neuron does not perform the integration process, so the latter should be taking place in another neuron of the network.

The observations above for a2, combined with the fact that this group in the *deaf* setup does not display antiphototaxis in *Env B*, suggest that the sound signalling system is connected to the discrimination between the two environments. In other words, the antiphototaxis is a result of the perception of the sound emitted by either robot. Furthermore, looking at Fig. 5d we observe that for group a2 (continuous line), Robot 2 leaves the band the moment Robot 1 emits a signal, despite the fact that its own sound output is not yet over the threshold of 0.5. We can summarise what happens as follows: the agent that "realises" first that its group has been placed in *Env B*, emits a sound signal, the perception of which triggers antiphototaxis in both robots of the group. We refer to this process as *external action selection*, since the selection of the appropriated action (i.e., the switch from phototaxis to antiphototaxis) is driven by the perception of an environmental cue (i.e., the sound signal) produced by either robot of the group. On the contrary, looking at the behaviour of group a10, we observe a process that we refer to as *internal action selection*, since the antiphototaxis is not triggered by a distinctive perceptual cue but solely by the internal dynamics of the neural network controller. While the *internal action selection* does not involve any form of communication, the *external action selection* determines the emergence of a simple form of communication between the robots of a group, since the robot that does not emit the signal initiates antiphototaxis by reacting to the other robot's signal.

The results of a pairwise Wilcoxon test among the fitness values of successful groups as recorded during 1000 evaluations in the *normal* setup, show that groups relying on an *external action selection* process to discriminate between *Env A*

and *Env B* (i.e., groups a1, a2, a7, a19) outperform with a confidence level of 99% the only successful group which relies on an *internal action selection* process (a10). We also compared the fitness scores achieved by the former groups in the *normal* setup, with fitness scores achieved if the communication channel between the robots is disabled, that is the robots are only capable of perceiving their own signals. The results show that these genotypes perform worse with a confidence level of 99% with the communication channel disabled with respect to the *normal* setup. These analyses seem to suggest that, once evolved through random mutations, mechanisms involved in the process of *external action selection* give to a group a selective advantage over those groups which do not possess these mechanisms.

This advantage might be related to the communication that results from the exploitation of an *external action selection* process. That is, by communicating the outcome of their decision about the state of the environment, robots may counterbalance the disruptive effect of the sensors and actuators' noise on the decision making mechanisms. In other words, the effectiveness of the mechanisms which integrate sensory information over time in order to disambiguate *Env A* from *Env B* may be sensibly disrupted by the noise inherent in the sensors' reading and in the outcome of any "planned" action. Equally, by communicating their decision, robots can eradicate decision delay between them: due to initialisation noise, one robot will on average perform the discrimination first. If the antiphototaxis is triggered by the perception of sound (*external action selection*) rather than by an internal state of the controller (*internal action selection*), then a robot which by itself is not capable or not yet ready to make a decision concerning the nature of the environment can rely on the decision taken by the other robot of the group. The former simply reacts to the sound signal emitted by the latter by initiating an antiphototactic behaviour (as happens for Robot 2 in Fig. 5d).

In order to test the "communication-noise hypothesis" (introduced above) as the main factor which determines the selective advantage of signalling over non-signalling groups, we run another set of twenty evolutionary runs. In these further series of simulations we removed any source of environmental noise which may interfere with the mechanisms for integration over time—i.e., no noise in sensors/actuators—and initialised the robots at exactly the anti-diametrical positions. Due to the latter choices, both robot controllers are at identical states during their lifetime and thus, a potential advantage related to the communication that results from the exploitation of an *external action selection* process is eradicated. The results of the post-evaluation phase are shown in Fig. 6. Groups controlled by the best genotypes of the last generation of runs 1 to 20 are called n1-n20, respectively. The results show that eleven of the groups effectively solved the task (n3, n4, n5, n6, n8, n9, n12, n14, n17, n18, n20). Table 2 shows results of the post-evaluation tests for two of the successful groups

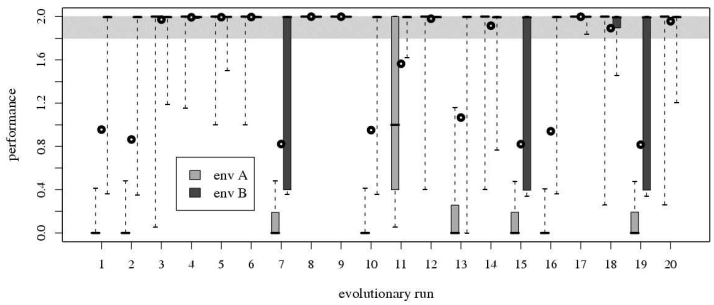

Fig. 6. Box-and-whisker plot visualising the post-evaluated fitness of groups n1-n20 in both environments. The box comprises observations ranging from the first to the third quartile. The median is indicated by a horizontal bar. When the observations are too close the box degenerates to the median. The whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range. The gray area denotes the area into which the average fitness value for both environments (black circles) must be, in order for the group to be called successful.

(n8, n17)[3]. These groups use signalling in *Env B* (see column 8), and given their failure to produce antiphototaxis in the *deaf* setup in *Env B* (see column 12), we can conclude that they employ an *external action selection* process to discriminate between *Env A* and *Env B* (exactly as a2). This suggests that there may be other factor(s) which cause the evolution of groups that exploit

[3] Other groups employ an *external action selection* process but use different signalling conventions, and others employ an *internal action selection* process. The analysis of these controllers is beyond the scope of this paper.

Table 2. Further results of post-evaluation tests with *normal* and *deaf* setup for two of the successful groups among n1-n20. For the *normal* setup, the table shows: (i) the average and standard deviation of the fitness over 500 trials in *Env A* (see columns 2, and 3) and in *Env B* (see columns 4, and 5); (ii) the average and standard deviation of the percentage of timesteps the sound was on by either robot over 500 trials in *Env A* (see columns 6, and 7) and in *Env B* (see columns 8, and 9). For the *deaf* setup the table shows the average and standard deviation of the fitness over 500 trials in *Env A* (see columns 10, and 11) and in *Env B* (see columns 12, and 13).

group	*normal*								*deaf*			
	fitness				*signalling (%)*				*fitness*			
	Env A		*Env B*		*Env A*		*Env B*		*Env A*		*Env B*	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
n8	2	0	1.999	0.000	0	0	26.40	0.64	2	0	1.169	0.007
n17	2	0	1.998	0.007	0	0	26.70	1.14	2	0	1.084	0.002

the *external action selection*. Genetic drift might be a possible explanation. Alternatively, we may speculate that sound evolves simply because, if emitted at the end of a complete tour around the light, it is a perceptual cue which an emitter robot can employ to initiate antiphototaxis. Robots that do not emit sound may find it more difficult to switch from phototaxis to antiphototaxis in the absence of a clear perceptual cue which triggers the latter response. Obviously, the latter hypothesis does not rule out the possibility that sound could acquire a communicative function in a subsequent time owing to its potential beneficial effect against environmental noise and decision delay on behalf of one robot, as explained above.

7 Porting on real robots

The task described in this paper is characterised by the fact that not only the change but also the persistence of particular sensorial states are directly linked to the effectiveness of the evolved strategies (see previous section). However, the evolved strategies are generated by robot controllers developed in a simulated world, which is responsible for modelling the sensory states of *s-bots* acting in *Env A* or *Env B*. If the physics of our simulated world are insufficiently and/or incorrectly defined, the evolved behavioural strategies may exploit loop-holes which strongly limit their effectiveness to an unrealistic scenario. Porting the controllers evolved in simulation onto a real robot is the ultimate proof to rule out the possibility of existence of the above mentioned circumstances (see [13]). However, as already pointed out in Section 1, this practice has not been taken into account in previous research work in which continuous time neural network controllers have been evolved to deal with tasks that required integration over time of sensory states. In this paper, we provide evidence of the "portability" of the evolved controllers, by showing the results of tests in which real robots of group a2 are repeatedly evaluated in *Env A* and *Env B*.

In [14], the author claims that the robot does not have to move identically in simulation and reality in order for the porting to be called successful, but its behaviour has to satisfy some criteria defined by the experimenter. Following this principle, real robots are considered successful if they carry out the main requirements of our task. That is, the robots have to reach the band in shades of grey regardless of the type of environment and subsequently (i) end up in the target area in *Env A*, without crossing the inner black edge of the circular band; (ii) end up as far as possible from the light in *Env B*. The robots should also avoid collisions.

Two *s-bots* (*s-bot*$_1$ and *s-bot*$_2$) were randomly positioned at a distance of 85 cm from the light, and 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*. We performed 40 trials, 20 in each environment. The results were 100% successful: there never was any

misdiscrimination, collision or crossing the black edge of the band[4]. As it was for the simulated robots of group a2, the *s-bots* accomplished the task through an *external action selection* process. That is, it is the sound emitted by one *s-bot* that triggers antiphototaxis in both robots. The results of our tests show that in *Env B* it is always *s-bot*$_1$ that emits a signal. Since the discrimination of an *Env B* from an *Env A*—that is, the emission of a sound signal and the following antiphototactic response—is based on the persistence of a particular sensory state, we can attribute the fact that *s-bot*$_1$ always signals earlier than *s-bot*$_2$ to mechanical and/or sensor differences between the two *s-bots*. However, the fact that we did not have any misdiscrimination proves that the simulation used to develop our controllers is sufficiently and correctly defined.

In order to understand to what extent real world noise influences the accomplishment of the task we performed further analysis. We compute the offset between the entrance position in the circular band of the robot that first emits a signal and the position at which this robot starts to signal. This measure, called offset Δ, takes value 0° if the robot signals exactly after covering a complete loop around the circular band. 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 emits a tone after having performed a loop around the light (see [4] for details on how to calculate the value of Δ).

As shown in Table 3, we see that the *s-bot* that first emits a signal—which, as mentioned above, is always *s-bot*$_1$—does so on average before completing a loop. However, being the magnitude of the offset Δ smaller than the width of the *way in* zone the group does not run into the risk of misinterpreting an *Env A* for an *Env B*. Further tests have proved that, if let to act alone in an *Env B*, *s-bot*$_2$ always signals after completing a loop (i.e., positive offset Δ, data not shown). This result can be accounted for by calling upon the inter *s-bot* differences, that can hardly be captured by the simulated world.

In Table 3 it is possible to compare the average offset Δ of a group of *s-bots* with the average offset Δ recorded by simulated robots of group a2. Contrary to the *s-bots*, the simulated robots signal on average after completing the loop. This result can be accounted for by calling upon the differences between simulated and real world. Quantitative descriptions of the behaviour of simulated and real robots can be definitely employed to evaluate the reliability of a simulated world as a tool to develop controllers for real robots. Our simulated world, which only

[4] The movies that correspond to all experiments can be found in `http://iridia.ulb.ac.be/supp/IridiaSupp2006-004`

Table 3. Average and standard deviation of the offset Δ over: (i) 20 trials in *Env B* performed by the *s-bots*; (ii) 500 trials in *Env B* performed by the simulated robots.

Offset Δ	avg	sd
s-bots	-30.6	11.75
simulated robots	+31.6	16.05

models a small subset of the *s-bot*-world physics, speeds up a particularly long evaluation process. Random noise has been employed as a tool to compensate for those physical phenomena not modelled (e.g., acceleration, friction, etc.). The results of our tests proved that the noise injected into the simulated world was sufficient to capture the variability of the behaviour of sensors and actuators of real hardware, which can easily disrupt the effectiveness of the evolved neural mechanisms. Of course, it should be mentioned that the simulated robots of group a2 are rather "conservative", that is, the robots emit signals rather late. This fact leaves a big margin of fault tolerance on which our successful porting to reality was also based.

8 Conclusions

In this work, we used artificial evolution as a means to engineer the emergence of communication in a group of robots but also to design robot controllers that can successfully cross the simulation-reality gap. Signals serve as "cues" that trigger behavioural switches in the group. Obviously, the evolved signalling system is simple, mainly due to the fact that we only allow agents to emit binary signals. In order to move to more complex signalling behaviours, we need to consider a sound system with more degrees of freedom, always in close relation to the task under consideration. Still, we observed that when agents are provided with "mouth" and "ears", communication can emerge, even without explicit fitness reward, providing groups that use it with a selective advantage over those that do not. Given this result, the question that arises is if we should aim at evolving communication in any swarm robotics task. Obviously this work does not provide enough evidence to answer positively. Any communication system that escapes from the local and simple interactions (e.g., communication through infra-red sensors—see [6]) might present disadvantages as well as advantages. In fact, when we move from a robot-to-robot to a robot-to-many interaction, not only the benefit of the knowledge of the environment acquired, but also possible errors spread faster. For example, in the task we studied, if a robot emits a signal, both robots can exploit it. However, if the signal is the product of a wrong decision (misinterpretation of environments) then both robots fail and the whole system collapses. The importance and the effect of such an event on the group performance is amplified as the swarm size increases. Therefore, issues like the reliability of signals have to be considered, together with the possible increase in the agent hardware and/or controller complexity that such, more advanced, forms of communication demand. To summarize, the experimenter has to balance the costs and benefits of communication before considering it as a path that might lead to the solution of a given task.

9 Acknowledgements

E. Tuci and M. Dorigo acknowledge European Commission support via the *ECAgents* project, funded by the Future and Emerging Technologies programme

(grant IST-1940), and by COMP2SYS, a Marie Curie Early Stage Training Site (grant MEST-CT-2004505079). The authors thank their colleagues at IRIDIA for stimulating discussions and feedback during the preparation of this paper, and the three anonymous reviewers for their helpful comments. M. 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. Bonabeau, E., Dorigo, M., Theraulaz, G.: Swarm Intelligence: From Natural to Artificial Systems. Oxford University Press, New York, NY (1999)
2. Trianni, V., Dorigo, M.: Self-organisation and communication in groups of simulated and physical robots. Biological Cybernetics (2006) In press.
3. Nolfi, S., Marocco, D.: Evolving robots able to integrate sensory-motor information over time. Theory in Biosciences **120** (2001) 287–310
4. Tuci, E., Trianni, V., Dorigo, M.: 'Feeling' the flow of time through sensory/motor coordination. Connection Science **16** (2004) 1–24
5. Nolfi, S., Floreano, D.: Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines. MIT Press, Cambridge, MA (2000)
6. Quinn, M., Smith, L., Mayley, G., Husbands, P.: Evolving controllers for a homogeneous system of physical robots: Structured cooperation with minimal sensors. Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences **361** (2003) 2321–2344
7. Nolfi, S.: Emergence of communication in embodied agents: Co-adapting communicative and non-communicative behaviours. Connection Science **17** (2005) 231–248
8. Mondada, F., Pettinaro, G.C., Guignard, A., Kwee, I.V., Floreano, D., Deneubourg, J.L., Nolfi, S., Gambardella, L.M., Dorigo, M.: SWARM-BOT: A new distributed robotic concept. Autonomous Robots **17** (2004) 193–221
9. Dudek, G., Jenkin, M.: Computational Principles of Mobile Robotics. Cambridge University Press, Cambridge, UK (2000)
10. Beer, R., Gallagher, J.: Evolving Dynamical Neural Networks for Adaptive Behavior. Adaptive Behavior **1** (1992) 91–122
11. Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, Reading, MA (1989)
12. Ampatzis, C., Tuci, E., Trianni, V., Dorigo, M.: Evolving communicating agents that integrate information over time: a real robot experiment. Technical Report TR/IRIDIA/2005-012, IRIDIA, Université Libre de Bruxelles (2005) This paper is available at `http://iridia.ulb.ac.be/IridiaTrSeries`.
13. Brooks, R.: Artificial life and real robots. In Varela, F., Bourgine, P., eds.: Towards a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life, MIT Press, Cambridge, MA (1992) 3–10
14. Jakobi, N.: Evolutionary robotics and the radical envelope of noise hypothesis. Adaptive Behavior **6** (1997) 325–368