On the coupling between agent internal and agent/environmental dynamics: Development of spatial representations in evolving autonomous robots

O. Gigliotta[12], S. Nolfi[1]
[1]Institute of Cognitive Sciences and Technologies, CNR, Viale San Martino della Battaglia 44, 00185, Roma, Italy
[2]Department of Psychology, University of Palermo, Viale delle Scienze, Ed.15, 90100, Palermo

Abstract

In this paper we describe how a population of evolving robots can autonomously develop forms of spatial representation which allow them to self-localize and to discriminate different locations of their environment by integrating sensory-motor information over time. Evolving robots also display a remarkable ability to generalize their skill in new environmental conditions never experienced before. The analysis of the obtained results indicates that evolved robots come up with simple and robust solutions that exploit quasi-periodic limit cycle dynamics emerging from the coupling between the robot/environmental dynamics and the robots' internal dynamics. More specifically, the variation of robots' internal states are governed by transient dynamical processes originating from the fact that robots' internal states tend to slowly approximate fixed attractors points corresponding to different type of sensory states which last for limited time duration and alternate while the robot moves in the environment.

Corresponding author: Stefano Nolfi, stefano.nolfi@istc.cnr.it, voice: ++39-06-44595233, fax: ++39-06-44595243

Keywords: spatial representation, evolution, transient dynamic.

1. Introduction

Recent advances in cognitive science (Varela et al., 1991; Brooks, 1991; Beer, 1995; Chiel and Beer, 1997; Clark, 1997; Pfeifer and Scheier, 1999; Nolfi and Floreano, 2000; Nolfi, 2005a) have clarified that behavior arises from the ongoing interaction between agents' nervous system, body, and environment. This implies that behavior and behavioral properties cannot be traced back to any one of these components in isolation. Indeed, agent/environmental systems which are simple from the point of view of their constituent elements (control system, body, and environment) might display complex behavioral and cognitive skills originating from the interaction between the constituent elements. Moreover, in agent/environmental systems in which the constituent elements have their own internal dynamics, complex behavioral and cognitive skills might originate from the interaction and the coupling of the dynamics occurring within the constituent elements or from their interaction.

In this paper, we demonstrate how a complex cognitive skill, which consists in the ability to discriminate different environmental locations and to self-localize, can emerge from the interaction between

two coupled dynamical processes (constituted by the robot's internal dynamics and the robot/environmental dynamics governed by simple rules). This will be illustrated in the context of a mobile robot situated in a maze environment which has to discriminate different spatial locations so as to recognize previously visited areas. The characteristics of the robot, of the environment, and of the task prevent the possibility to solve the problem on the basis of the information available to the robot in any single moment and require an ability to recognize abstract features of the agent/environmental interaction by integrating sensory-motor information through time (Beer, 2003; Nolfi & Marocco, 2001; Slocum et al., 2000; Nolfi, 2005b).

The problem of discriminating different spatial locations is a key feature of animal navigation. Indeed, although spatial recognition skills are not necessarily required to exhibit simple navigation forms, they are certainly crucial to solve more complex navigation problems (Trullier et al, 1997). However, this research also aims to study, more generally, perceptual categorization problems in which an agent should discover abstract categories, identify the categories that fit the current agent/environmental situation by integrating sensory-motor information over time, and use the outcome of the perceptual categorization process to act appropriately. More specifically, we aim to investigate perceptual categorization processes from an active perception perspective (Beer, 2003; Morimoto & Ikegami, 2004; Nolfi 2005b) in which the way in which sensory information change over time is co-determined by the way in which the agent reacts to the perceived sensory states.

The methodology that we will use to achieve these goals is synthetic (i.e. is based on the attempt to synthesize the behavioral and cognitive skills that we want to study in artificial robotic agents) and is based on a self-organizing techniques in which important characteristics of the robots are not designed by the experimenter but are developed autonomously by the robots themselves while they try to adapt to the environment and to the problem they have solve. The reason for using a self-organizing technique is due to the fact that we are interested in studying whether robots can develop and use an effective categorization abilities autonomously rather than developing robots that can solve their task on the basis of hand-crafted solutions. We chose to use artificial evolution, rather than other learning techniques, since it leaves the robots free to determine the way in which their problem should be solved (i.e. since it allows limiting the constraints on how the problem is solved to the minimum). However, we expect that similar results can be obtained by using other learning techniques provided that: (a) the free parameters that are varied during learning regulate the fine-grained interactions between the robot and the environment, and (b) variations are retained or discarded on the basis of their effects on the overall behaviour exhibited by the robot (Nolfi, 2005a).

The results described in this paper also provide a contribution for what concern the question of how robots can represent spatial information (e.g. the position of selected locations of the environment) and more generally to the question of how representation can self-organize in embodied and situated agents which adapt to their task/environment. The former aspect will be briefly discussed in the concluding section of the paper. The discussion of the latter aspect, instead, goes beyond the goals of this paper (but see Beer, 2003). Since the term representation is often used with different meanings, it is important to specify that in the rest of the paper we will use this term to indicate agents' internal states which co-vary with agents' positions and orientations in physical space.

2. Related Literature

In this section, we review the related literature by restricting our analysis to research works in which the way in which robots encode spatial information is not pre-determined by the experimenter and is developed by the robots themselves through an adaptive process while they interact with the environment.

In a recent work, Vickerstaff and Di Paolo (2005a, 2005b) evolved simulated robots (provided with a compass sensor, speed sensors, and two light sensors) for exhibiting a homing behaviour. The robots, which are initially placed in their home location, are selected for the ability to return to the same location after

having reached a variable number of light beacons placed randomly. Beacons are placed one at a time --- a new beacon is displaced as soon as the robot reaches the previous beacon. Robots were provided with a CTRNN controller (Beer, 1995a). The free parameters, which were encoded in evolving genotypes, encoded the connection weights, the time constant of the neurons, and the architecture of the neural controller. The analysis of the obtained behaviour indicates that evolved robots solve their problem by using a path integration mechanism, i.e. by encoding the homing location in the activation state of a vector of neurons (or a vector of modifiable weights, in a variation of the experiment). This internal vector is continually updated by integrating sensor information over time while the robot travels in the environment so as to always encode an updated estimation of the current position of the robot relative to the home location. Evolved robots also display a searching behaviour, once they have reached their estimate of the nest location, similar to that exhibited by ants.

As in the case of Vickerstaff and Di Paolo's work, we develop robots able to display a navigation ability by leaving them free to determine the way in which they internally encode spatial information. The main difference with the work described in this paper lays in the nature of the problem. The work of Vickerstaff and Di Paolo involves a homing problem that consists in returning to a recently visited environmental location and that can be solved through a path integration method --- a navigation method similar to that employed by sailors under the name of dead reckoning that consists in continuously updating the estimation of the relative location of the reference point (the nest) on the basis of the estimated direction and speed of the movements performed after abandoning the reference point. The work described in this paper instead involve a place recognition problem (Trullier et al, 1997) in which robots should be able to discriminate their current location by identifying regularities in their sensory-motor flow while they move in the environment.

In an earlier related work, Floreano and Mondada (1996) evolved a Khepera robot (Mondada et al., 1993), provided with infrared, light, and battery level sensors, for the ability to move as straight and as fast as possible by periodically returning to a recharging area located close to a light beacon to recharge its battery. Evolving robots display an ability to integrate spatial and energy consumption information so as to allow the robot to periodically return to the charging station just before the energy level goes below a "survival" threshold. Since the light gradient provided a straightforward indication of the recharging area and it is visible from most of the environmental locations, evolving robots do not need to evolve a path integration mechanism to identify the relative location of the target. Moreover, the relative location of the target can be identified in any single time step on the basis of the current state of the light sensors. The main difference between Floreano and Mondada and our experimental scenario is that, in our case, robots cannot infer their location in the environment on the basis of a single sensory pattern only since robots experience identical sensory patterns in different environmental locations.

Tani and Fukumura (1997) have demonstrated how a mobile robot provided with a recurrent neural network controller trained through a supervised learning algorithm can develop an ability to navigate in a maze environment and to select the right route in brunching points so as to produce the required periodic behaviour (i.e. a repeated 0 or 8 shape trajectory). As in the case of the experiments described in this paper, after few repetitions of the required trajectory the internal states of the trained neural controller converge toward a limit cycle dynamic which allows the robot to self-localize so as to select the right route at branching points. One differences between this work and the experiments described below concerns the fact that in our model the regularities which drive the robot's internal dynamics are selected by the robots themselves and are not defined by the experimenter. A second difference concerns the characteristics of the task and of the robots' internal dynamics which, in the latter case, allow the robot to generalize their self-localization skills in different environments eliciting different periodic behaviours.

Nolfi (2002) evolved a simulated Khepera robot, provided with infrared sensors, to navigate into a two room environment by discriminating the room in which it is currently located. As in the case of the experiment described in this paper, robots cannot rely on a path integration method since robots' initial

position in the environment is randomly initialised. Moreover, as in the case of the work described in this paper, robots experience identical sensory states in different environmental locations, and therefore should be able to discriminate their current environmental location by integrating sensory-motor information over time. The work described in this paper is related to this previous work but present new features that allow robots to display a greater expressive power and remarkable generalization abilities.

Finally, in two related set of experiments Tani & Nolfi (Nolfi & Tani, 1999; Tani & Nolfi, 1999), investigated how a wheeled robots provided with distance-proximity sensors travelling in an multi-room environment could develop an ability to predict the forthcoming sensory states by segmenting sequences of sensory states into categories and by anticipating the category of the forthcoming sensory states. As in the case of the experiments described in this paper, in the works referred above and in other related researches (Wolpert and Kawato, 1998) the categories which allow the robots to recognize different place of the environments are not designed by the experimenter but are discovered by the robots while they learn to solve their problem in interaction with the environment. The differences with respect to the work presented in this paper concern the fact that in this case: (a) robots are not forced to use discrete categories, and (b) robots are not requested to anticipate future events but only to self-localize and to recognize previously visited locations.

3. The experimental setup

Consider the case of an E-puck robot (Figure 1, left) placed in a double T-maze environment (Figure 1, right) that should be able to explore the environment, memorize the location of the target area, and recognize that location after being moved to a new position in the environment. The location of the target area is marked by a black disk placed on the ground which can be easily detected by the robot through an infrared sensor pointing toward the ground (Figure 2, centre). The black disk, however, is removed after the target area is reached by the robot for the first time (Figure 2, right). To recognize the location of the target area after the disk has been removed, therefore, the robot should keep in memory an indication of the estimated position of the target area and recognize when it reaches the same location again.

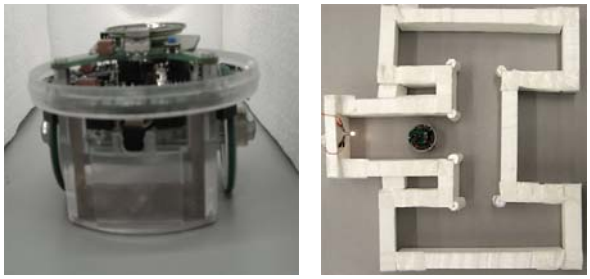


Figure 1. Left: The e-puck robot developed at EPFL, Switzerland (http://www.e-puck.org/). Right: The double T-maze environment including a light bulb located at the end of the central corridor.

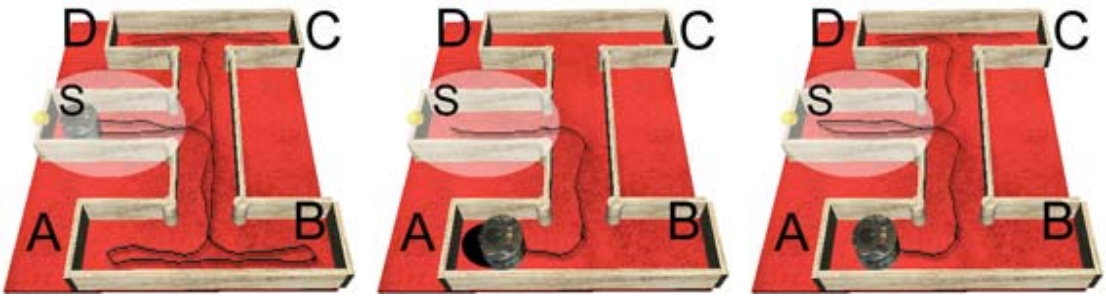


Figure 2. The robot and the environment in simulation. The white circle indicates the amplitude of the light gradient. The black line indicates the trajectory produced by a typical evolved robot navigating in the environment. **Left**: At the beginning of a trial the environment does not contain the black disk. **Center**: At a certain point a black disk is placed in one of the four possible locations indicated with letters A, B, C, and D. **Right**: The black disk is removed and the robot is moved to location S or to a new randomly selected position in the environment (see text).

The T-maze arena covers an area of 52cm by 60cm. The robot has eight infrared sensors (which provide information about nearby obstacles up to a distance of about 5cm), eight light sensors (which provide the light gradient information up to a distance of about 30 cm from the light), one ground sensor (which detects the colour of the ground), and two motors (which control the desired speed of the two corresponding wheels).

Each robot is controlled by a neural network with a fixed architecture including 18 sensory neurons (which encode the state of the eight infrared sensors, of the eight light sensors, and of two location sensors which will be described below), two internal neurons, and three motor neurons. Two motor neurons encode the desired speed of the two wheels. One output neuron encodes whether the robot believes to have reached the previous location of the black disk or not. Sensory neurons therefore do not include information about the ground sensor. The role of this sensor, in fact, is that to decide when the current activation state of the two internal neurons should be recorded. The role of this sensor, in fact, is that to determine the state of the internal neurons which should be associated to the location of the target (see below).

The lifetime of individual robots consists of several trials involving two successive phases. At the beginning of the trial the robot is placed in position S with a randomly selected orientation and is allowed to move in the environment (Figure 2, left). At a certain point (i.e. after about 90 s) a black disk is placed in one of the four possible target locations indicated with the letters A, B, C, and D. The first phase is terminated as soon as the robot encounters the black disk (Figure 2, center). During the second phase the robot is placed to position S with a randomly selected orientation, the black circle is removed, and the robot is allowed to move for the remaining part of the trial. The second phase terminates as soon as recognition neuron is triggered on (i.e. as soon as the activation of the neuron exceed a threshold of 0.5) or after 240 s (330 s from the beginning of the trial). The trial is considered successful or unsuccessful depending on whether the robot triggers on its recognition neuron when it reaches the previous location of the black disk or not. During the first phase of each trial in which the robot has not yet reached the black disk, the activation of the location sensors is set to the maximum value (1.0). During the second phase of each trial, the state of the two location sensors is set to the absolute difference between the current activation state of the internal neurons and the state of the internal neurons previously associated to the position of the target.

Output neurons are updated according to the logistic function (see equation 1). Internal neurons are updated according to equation 2 (see also Nolfi, 2002).

$$O_output_j = \left(1 + e^{-A_j}\right)^{-1} \tag{1}$$

$$O_internal_j = \tau_j O_internal_j^{(t-1)} + \left(1 - \tau_j\right)\left(1 + e^{A_j}\right) \tag{2}$$

With $Oj^{(t-1)}$ being the output of the *j*th neuron at the previous time step, *Aj* being the activity of the *j*th neuron, τj the time constant of the *j*th neuron. The activity of internal and motor neurons is computed according to the following function:

$$A_j = t_j + \sum w_{ij} O_i \qquad (3)$$

With *tj* being the bias of the *j*th neuron, *Wij* the weight from the *i*th to the *j*th neuron, *Oi* the output of the *i*th neuron that sends a connection to the *j*th neuron. The state of the neurons is updated each 100 ms.

The important aspect to note here is that internal neurons consist of leaky integrators neurons, i.e. neurons which hold a certain amount of their activation from the previous time step and in which the effect of the previous state of on their current state is determined by their time constant parameter (for alternative ways to implement leaky neurons see, for example, Beer [1995b]). These neurons thus are dynamical systems operating at tuneable time scales. The time constant parameter, in fact, regulates how quickly or how slowly the output of the neurons varies in time. As we will see, the dynamical system nature of the internal neurons and the possibility to set the time scale at which they operate is a crucial pre-requisite for the solutions found by evolving robots.

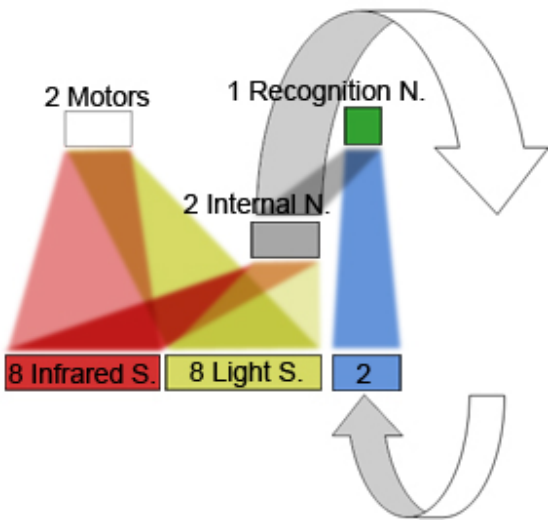


Figure 3. The architecture of the robot's neural controller. The grey areas indicate the connections between blocks of neurons. The arrow indicates that, during the second phase, the activation of the two additional sensory neurons encodes the absolute difference between the current state of the two internal neurons and their previous state recorded when the robot detected the black disk. The recognition output neuron should be turned on when the robot re-visits the location where the black disk was previously located.

The free parameters of the robots' neural controllers (i.e. the strengths of the connection and the time constant parameters of the two internal neurons) are evolved (Nolfi & Floreano, 2000). The initial population consisted of 100 randomly generated genotypes which encode the connection weights, the biases, and the time constants of the two internal neurons of 100 corresponding neural controllers. Each parameter is encoded with 8 bits and normalized in the range [–5.0, +5.0], in the case of connection weights and biases, and in the range [0.0, 1.0], in the case of time constants. The 20 best genotypes of each generation were allowed to reproduce by generating five copies each, with 2% of their bits replaced with a new randomly selected value. The evolutionary process lasted 800 generations (i.e. the process of testing, selecting and reproducing robots is iterated 800 times). The experiment was replicated 10 times.

Robots are evaluated for 32 trials lasting 3300 cycles, with each cycle lasting 100ms. Robots are evaluated with respect to their ability to explore the environment, and recognize the previous location of the

black disk in two successive periods consisting of 4 and 28 trials, respectively. However, the second period only occurs when robots reach a fitness equal or greater than 160 during the first period.

During the first period the environment is divided into 20 target areas equally distributed in the double T-maze and the robot is scored with 1 point for each newly visited target area. The fitness value is then truncated to a value of 160. This threshold has been set so as to encourage the robot to explore all areas of the environment (i.e. all the 20 different parts of the environment) at least 8 times during the first 4 trials without forcing the robot to maximize its speed as fast as possible. During the second period the robot is scored with 1 point for each trial in which it successfully recognizes the previous location of the black disk (after it has been removed and after the robot has been moved to location S).

By analysing the value of the fitness throughout generations for one of the best replications of the experiment (see Figure 4), we can identify an initial phase, lasting about 200 generations, in which the evolving robots develop an ability to navigate in the environment so as to visit all sub-areas followed by a phase lasting about 400 generations, in which the robots develop an ability to self-localize and to recognize previously visited locations of the environment.

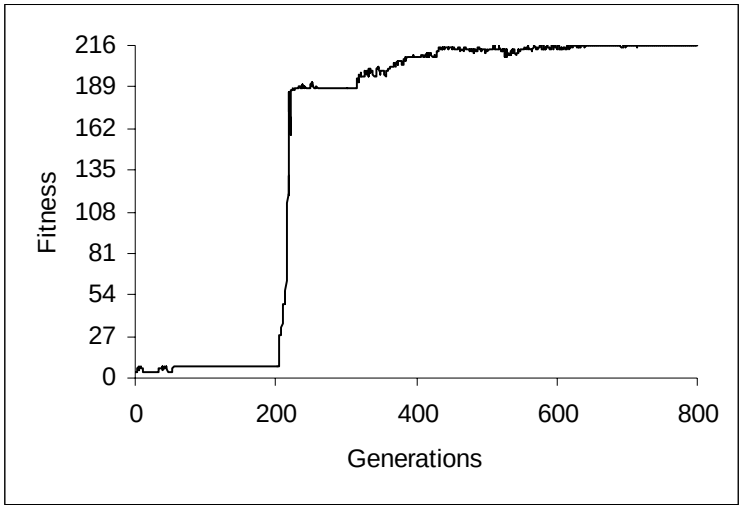


Figure 4. Fitness throughout generations for one of the best replications of the experiment. Since the maximum score for the exploration task is 160, and the maximum score for the recognition task is 56, optimal performance on the two tasks corresponds to a score of 216.

4. Results

By analysing the behaviour displayed by the best individual of the last generation for the best replication of the experiment we observed that the robot displays an effective navigation behaviour which allows it to periodically visit all the locations of the environment by moving forward in corridors, turning right on T-junctions, and turning back at corridors' end points (see Figure 5, left). Moreover, the evolved robot displays an optimal ability to recognize the previous location of the black disk by always turning its recognition neuron on in the location in which the black disk was previously located and by never turning the recognition neuron on in wrong locations (Figure 5, right, white histograms). Evolved robots also display an ability to generalize their ability when they are tested in a condition in which they are placed in a randomly selected location of the environment at the beginning of each trial and after the black disk has been removed (Figure 4, right, grey histograms). Finally, by downloading the controller evolved in simulation onto the robot and by testing the robot in the physical environment, we observed that the

performance obtained in the real environment are comparable with that obtained in simulation (Figure 4, right, black histograms).

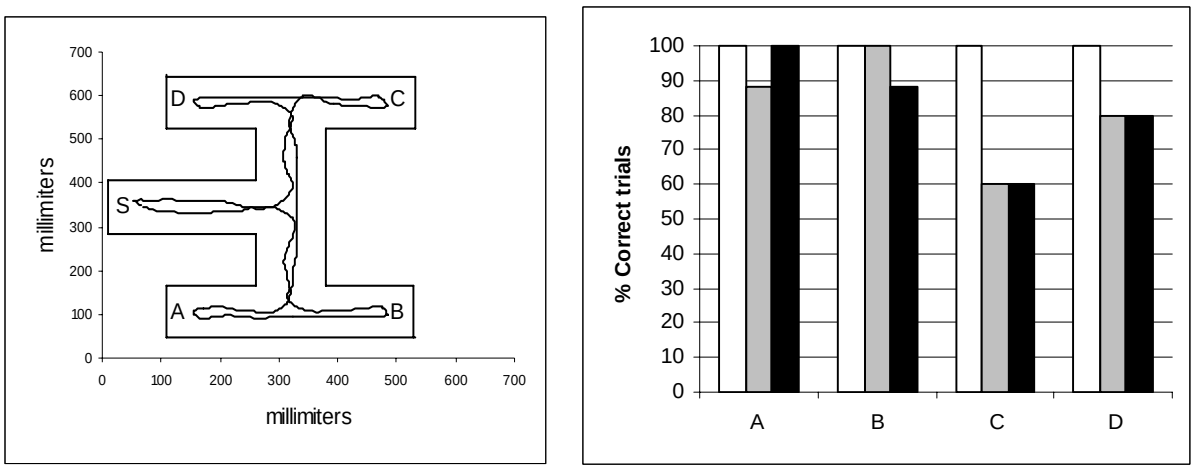


Figure 5**.** Motor behaviour and performance on the recognition task for the best robot of the best replication of the experiment. **Left**: The trajectory produced by the robot during one lap of the environment. **Right**: Percentage of trials in which the robot is able to recognize the previous location of the black disk. Data obtained by testing the robot for 100 trials in each condition. The four groups of histograms indicate the performance for the four possible location of the target area. White histograms indicate the performance obtained in simulation by moving the simulated robot to location S after the black disk has been detected (see Figure 2). Grey histograms indicate the performance obtained in simulation by moving the simulated robot to a randomly selected location after the black disk has been detected. Black histograms indicate the performance observed in hardware by moving the robot to a randomly selected location after the black disk has been detected

Evolved robots display an ability to generate internal states which are different when the robot is located in different parts of the environment and which are similar when the robot visits or re-visits the same environmental location, by integrating sensory information over time. Indeed, by plotting the state of the robots' internal neurons recorded when the robot is located in the four possible locations of the target area (Figure 6) we can see how the states corresponding to the same location are clustered together and the states corresponding to different locations are separated in the two dimensional state space which encodes all possible states of the internal neurons. These data have been obtained in the condition in which robots are moved to a randomly selected location after the black disk has been removed.

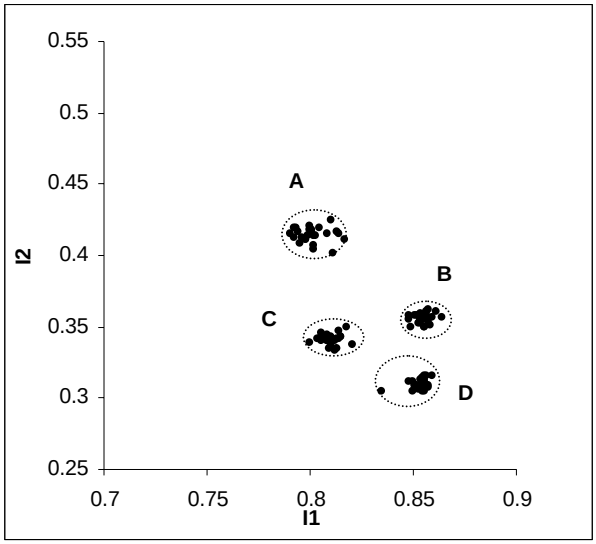


Figure 6**.** The black dots indicate the activation state of the internal neurons of the best evolved robot when it detects the black disk for the first time. Data correspond to 25 trials for each of the four possible positions of the target area for a total of 100 trials. The two axes (I1 and I2) indicate the activation state of the two corresponding internal neurons. The letters A, B, C, and D, and the circles indicate how activation states corresponding to similar locations in space are clustered in different areas of the state space of the two internal neurons.

4.1 Robots ability to represent environmental locations

In this section, we analyse how evolved robots represent different locations of the environment, how they generate their representation while they navigate in the environment, and how they are able to generalize their skills (to a certain extent) with respect to the position in which they are placed at the beginning of the trial and with respect to the position to which they are moved one they found the black disk. We are using the term representation to indicate a robot's internal state which co-varies with the robot's location in the environment or with the characteristics of the environment in which the robot is situated.

By analysing the dynamics of the two internal neurons while the robot moves in the environment we can see how after 95 s (i.e. after the robot makes about 2 laps of the environment) the state of the internal neurons converges toward a limit cycle attractor (Figure 7). By analysing the dynamics of the internal neurons when the robot is situated in the environment in this and in other replications of the experiment, we observed the following characteristics: (1) *periodicity*: the dynamics of the internal neurons produces a limit cycle which repeats itself every time the robot performs a lap of the environment, and (2) *expressiveness*: different points along the trajectory of the limit cycle correspond to different locations of the robot in the environment (beside few cases in which the same point correspond to different locations). For a description of how periodicity and expressiveness have been measured, see below.

As shown in Figure 7, the shape of the limit cycle significantly varies in different replications of the experiments. Indeed, periodicity and expressiveness might be realized through a wide variety of internal dynamics.

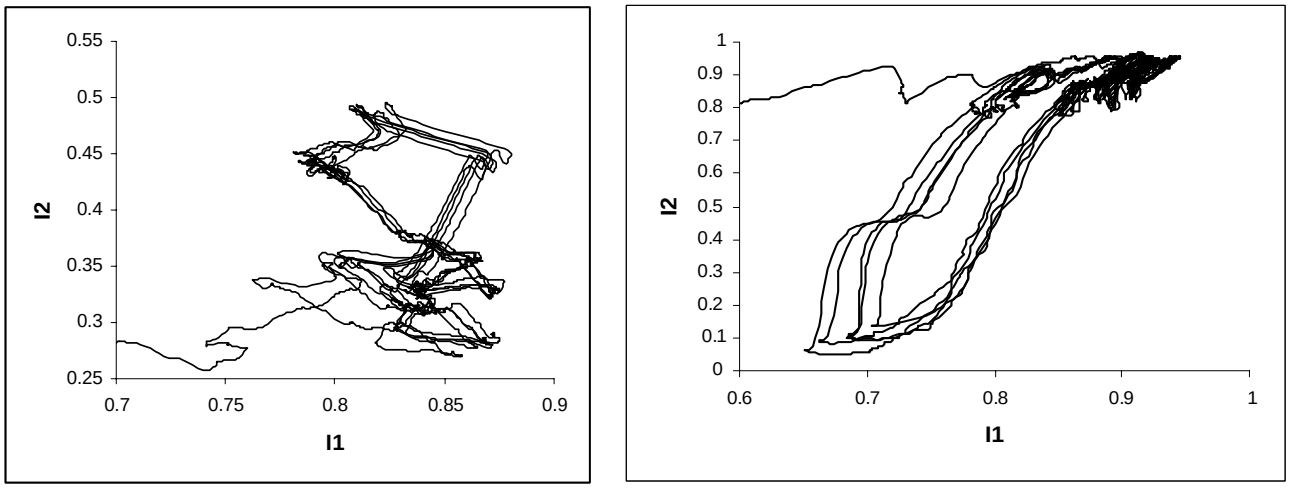


Figure 7. **T**he activation state of the internal neurons recorded while the robot moves in the environment for 330 s. **Left**: data for the best replication of the experiment. **Right**: data for the second best replication of the experiment.

The periodicity of the limit cycle explains how the evolved robots generalize their skill with respect to its initial position. In fact, after a certain amount of time in which the robot moves in the environment, the state of the robots' internal neurons correlates only with the sensory states experience during a relatively short time period but not anymore with the sensory states experienced before that period which in turn depends on the robot's initial position and orientation. The expressiveness of the limit cycle ensures that different locations of the environment correspond to different internal states. The combination of these two properties ensures that the robot can reliably represent its position in the environment so as to distinguish between location A, B, C, and D and so as to self-localize after having travelled in the environment for a sufficient amount of time independently from its initial position and orientation.

These periodic limit cycles result from a transient dynamic which originates from the coupling between the robot's internal dynamic and the robot/environmental dynamic. The robot's internal dynamics originate from the fact that the state of the two internal neurons tend to move toward different attractors points (Figure 8) corresponding to different type of sensory states with a speed which is proportional to the time constant parameters of the neurons. The robot/environmental dynamics originate from the fact that different type of sensory states, corresponding to different fixed point attractors last for different time durations and alternate while the robot moves in the environment. The coupling between these two dynamical process results from the fact that the parameters which regulate the robot's internal dynamics (i.e. the time constant of the leaky neurons and the connection weights which determine the position of the fixed point attractors in the neurons' state space) and the parameters which regulate the robot/environmental dynamics (i.e. the speed and the trajectory with which the robot move in the environment) are co-evolved. For related examples of how transient dynamics and coupling between robot's internal and external dynamics can be exploited in the context of embodied and situated agents see, Tani and Fukumura [1997] and Beer [2003]).

For example, when the robot is going to terminate one lap of the environment, the state of the internal neurons occupies the central position of the limit cycle (see Figure 9 "s"). Then, while the robot travels toward the light, the state of the internal neurons moves toward the light-forward (LF) average fixed point attractor situated in the top-right corner of the state space. Then, when the robot moves away from the light, its internal state tends to move toward the light-rear (LR) average attractor located in the top-left corner of the state space. Later on, when the robot turns right and then moves along the corridor, its internal state moves toward the no-obstacles (NO) and obstacle-right (OR) attractors located in the left and right side of

the state space respectively. These latter movements bring the state of the internal neurons in the area indicated with the letter A that is experienced when the robot reaches the corresponding location of the double T-maze environment (see Figure 2). The subsequent movements toward the fixed point attractors corresponding to the sensory states experienced later on produce the following path of the limit cycle until a new lap of the environment and a new lap of the corresponding limit cycle starts.

The fact that the fixed point attractors are never fully reached by the state of the two internal neurons while the robot navigates in the environment (since the corresponding sensory states are experienced only for a limited amount of time and since the state of the internal neurons move slowly toward the attractors) ensures that the states of the two internal neurons preserve information about previously experienced sensory states.

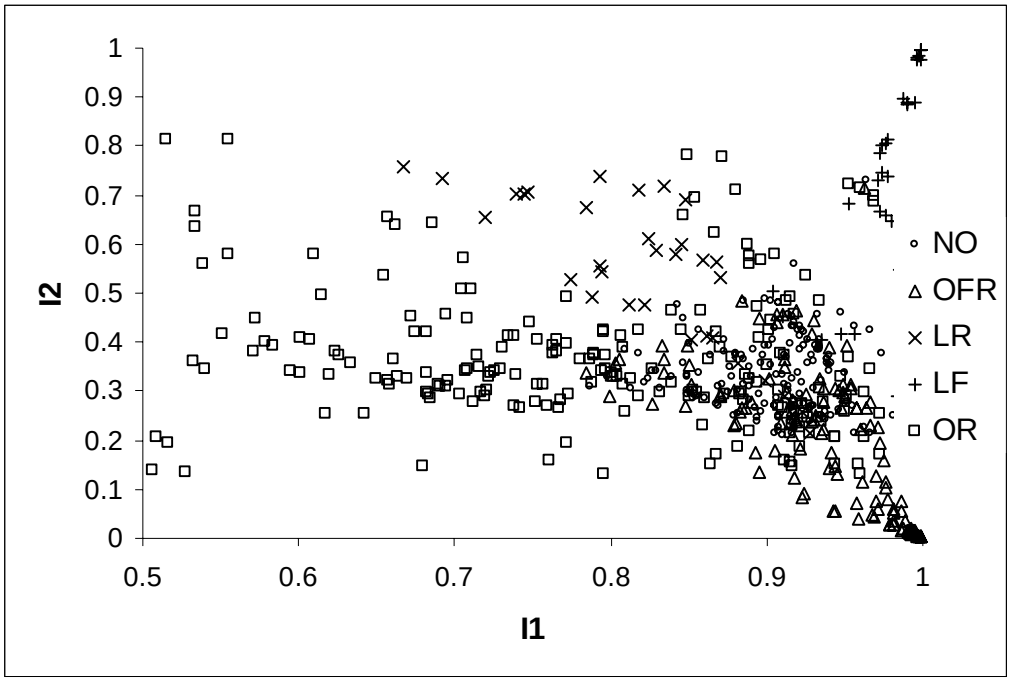


Figure 8. Fixed point attractors in the internal neurons' state space corresponding to 3300 sensory states experienced by the robot during one lap of the environment lasting 330s. These sensory states and the corresponding fixed point attractors have been manually subdivided into 5 different classes which correspond to qualitatively different types of sensory states and which are visualized with different labels. These classes include sensory states in which the robot detects: a light in its frontal side (LF), a light on its rear side (LR), an obstacle on its right and frontal side (OFR), an obstacle on its right side (OR), no obstacles and no lights (NO).

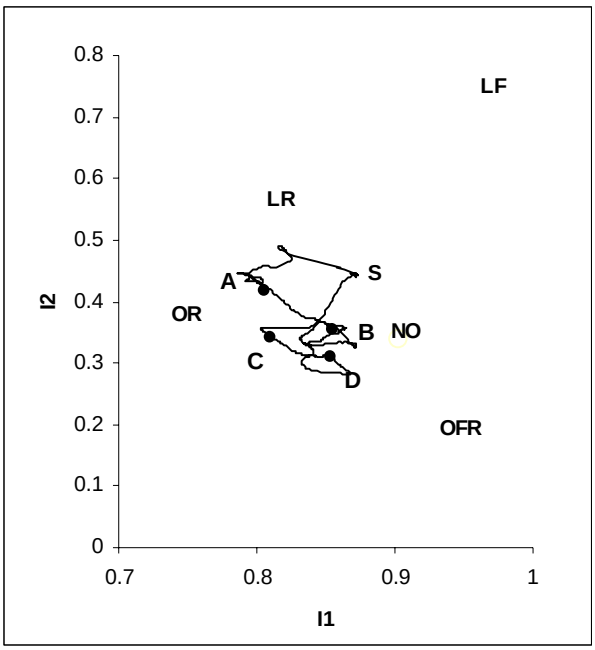


Figure 9. A schematization of the limit cycle dynamics shown in Figure 7, right. The labels indicate the internal state of the two neurons when the robot reach locations A, B, C, D, and S (Figure 2). The LR. LR, OR, OFR, and NO labels indicate the average position of the five classes of fixed point attractors shown in Figure 8.

Figure 10 and 11 show the level of expressiveness and periodicity of the internal representation of the best robot of successive generations in the case of the best replication of the experiment. The data to be analysed has been collected by placing the robot in a randomly selected position and orientation in the environment and by recording the Cartesian coordinates of the robot and the state of the two internal neurons every 100ms while the robot complete the last 3 out of 4 laps of the environment (i.e. we discard the data of the first lap during which the robot might not be self-localized yet).

The analysis was replicated in two conditions in which only the data corresponding to the four possible location of the black disk or all the data collected were analyzed, respectively. In the former condition, the data was divided into four categories corresponding to the four possible locations of the black disk, while in the latter condition the data was divided into 100 categories corresponding to 100 adjacent portions of space occupied by the robot during an entire lap of the environment.

Expressiveness, that is robots' ability to represent the same location with similar internal states and different locations with different internal states was measured by calculating the percentage of times in which a given item (i.e. a given state of the two internal neurons) and its most similar item belonged to the same category or not (for a similar index, see Thornton and Clark, 1997). Similarity between items was measured by calculating the absolute difference between the two corresponding vectors. A value of 1.0 thus indicates an optimal expressiveness value in which the internal state of the robot in a certain location of the environment is always more similar to the internal state of the robot in approximately the same location than to the state in other locations.

Periodicity, i.e. robots' ability to represent the same spatial location with the same state of the two internal neurons independently from the robot's initial position and orientation, has been measured by calculating the standard deviation of the state of the two internal neurons for items corresponding to the same location summed over all locations. A periodicity value of 0.0 thus corresponds to an optimal value in which, for all considered locations, the same space location is represented by identical internal states.

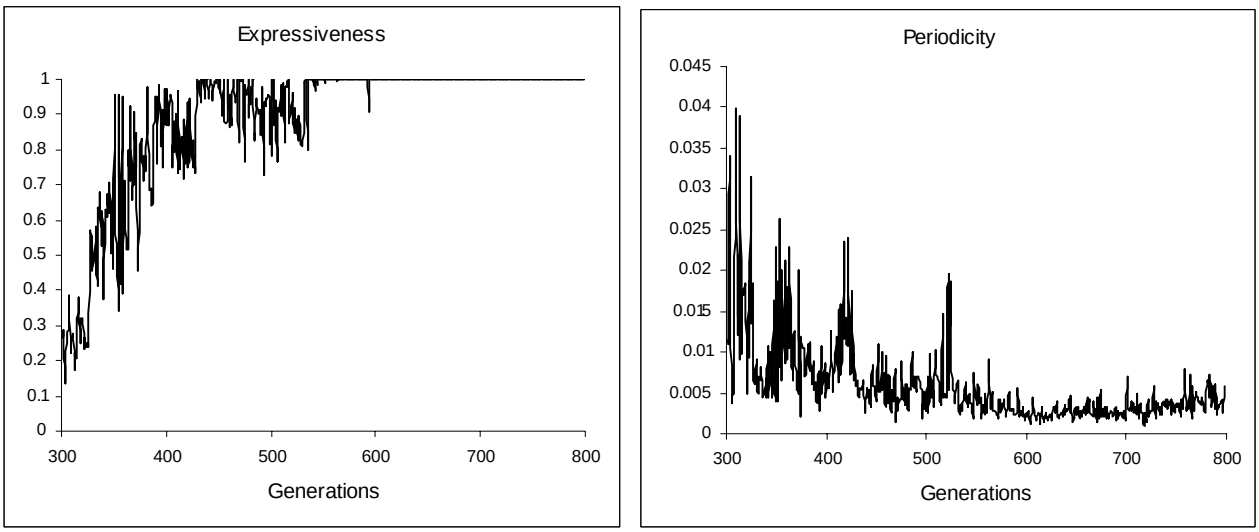


Figure 10. Expressiveness and periodicity (left and right figures) calculated on the 4 categories corresponding to the four possible locations of the black disk in the double T-Maze environment for the best robot of succeeding generations in the case of the best replication of the experiment.

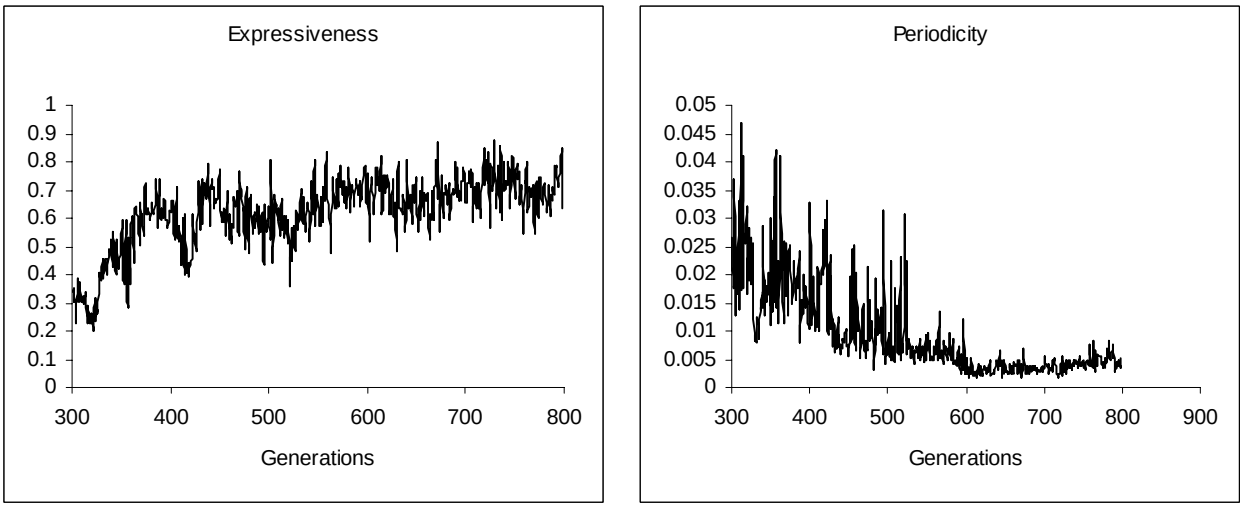


Figure 11. Expressiveness and periodicity calculated on 100 locations distributed uniformously within the double T-Maze environment for the best robot of succeeding generations for the best replication of the experiment.

The analysis of the expressiveness and periodicity for the four areas of the environment corresponding to the four possible locations of the black disk indicates that expressiveness reaches the maximum value during the first 550 generations (Figure 10, left) and that periodicity reaches a close to optimal value after 600 generations (Figure 10, right). The analysis of the expressiveness and periodicity for the 100 areas uniformously distributed in the environment indicates that expressiveness increases by reaching a value of 0.7 after about 700 generations and that periodicity improves by reaching a close to optimal value after about 600 generations (Figure 11).

The fact that expressiveness in the latter case is not fully optimized can be explained by considering that evolving robots are required to differentiate only the four areas of the environment which correspond to the four dead-ends of the two external corridors. The fact that the level of

expressiveness is much higher than that requested to differentiate the four possible locations of the black disk, can be explained as a generalization effect. In order to reliably differentiate the four areas, in fact, robots should be able to keep track of the characteristics of the environmental areas located near those four areas (i.e. to also differentiate the other areas located nearby). Indeed, thanks to this generalization effect, the evolved robot shows an ability to discriminate different portions of the environment in about 70% of the cases despite the fact that it has been evolved for the ability to discriminate a much smaller number of cases.

4.2 Robots' ability to represent different environmental topologies

In this section we analyse how robots generalize their abilities when tested in environments that have different topologies from the environment in which they have been evolved. As we will see, the combination of robots/environmental and robots' internal dynamics produce different limit cycles in different environments. The fact that the limit cycles produced in new environments tend to also show expressiveness and periodicity ensures that the robots are able to generalize their ability to self-localize and to recognize previously visited locations in environments which differ (within limits) from the environments in which they have been evolved without the need of any further training.

To verify the robots' generalization abilities and to analyse the dynamics of the internal neurons in different environmental conditions we tested the best individual of the best replication evolved in the double T-Maze environment in a simple T-maze environment, in a simplified double T-Maze environment, and in an open-arena environment (Figure 12). As shown in Figure 13, the robot displays a rather good ability to correctly recognize the previous location of the black disk in the new environments within two or three possible locations indicated with the letter A, B, and C. The limit cycles produced by the robot in the new environments are shown in Figure 14.

To understand why the shape of the limit cycle varies depending on the environmental topology we should consider that it is the result of two factors: (1) the position of the fixed point attractors in the internal neurons' state space, and (2) the sequence of sensory state experienced by the robot while it moves in the environment. The first factor is a function of the characteristics of the neural controller and therefore does not change when the robot is situated in a new environment. The second factor, however, obviously depends on the characteristics of the environment in which the robot is placed.

Overall, this implies that robots can generalize their abilities in a wide range of environments with different spatial organizations provided that the local characteristics of the new environments are similar to that of the environment in which they have been evolved (i.e. in this particular case, provided that the new environment has light gradients, corridors, and dead-ends) and provided that the size of the environment is comparable to that of the environment in which they have been evolved (i.e. provided that the time duration of perceptual events in the new environment is comparable to the time duration of events in the environment in which the robot has been evolved). The latter point can be explained by considering that the time constants that determine the rate at which the state of the internal neurons vary toward the corresponding fixed point attractors should be sufficiently small to avoid reaching the corresponding fixed point attractors and sufficiently big to produce enough variation in the state of the internal neurons while the robot moves in different locations of the environment. Indeed, by testing the evolved robot in a double-T maze environment in which the length of the corridors was doubled with respect to the environment used during the training process, we observed that the performance decrease from 100% to 49% of correct localizations (with a percentage of correct localization of 100, 56, 36, and 4 for target A, B, C, and D, respectively).

Interestingly, while the state of the internal neurons can be described as a representation of the robot's relative position in the environment, the limit cycles consisting of the sequences of internal states generated while the robot navigates in the environment, can be described as a representation of

the topology of the environment. In fact, different environments correspond to different limit cycles. Moreover, the level of similarity between two environments or between two sub-parts of two environments is reflected in the level of similarity of the corresponding limit cycles or of the corresponding sub-parts of the limit cycles (see Fig. 6 and 13).

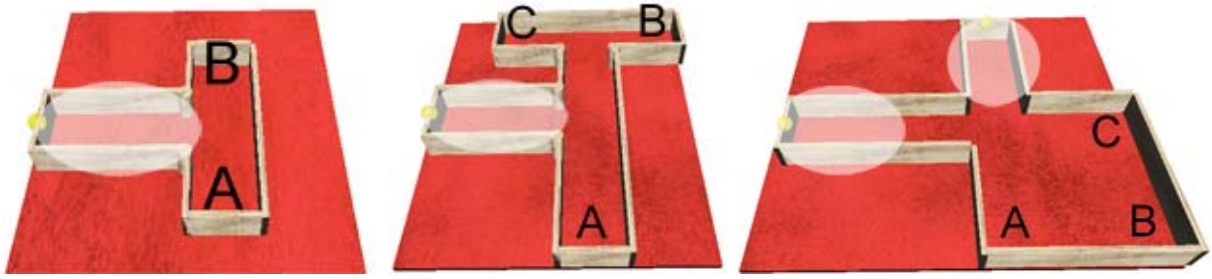


Figure 12**.** Left: The T-Maze environment. Centre: The simplified double T-maze environment. Right: an open arena environment with two light bulbs.

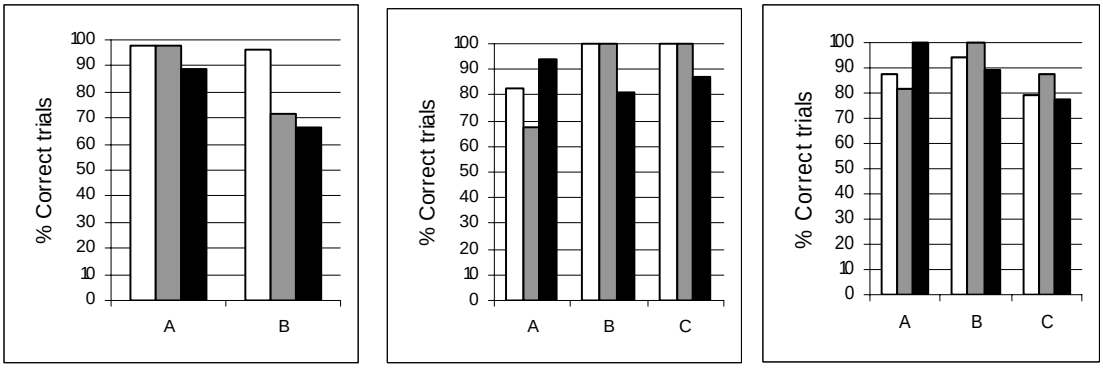


Figure 13**.** Percentage of trials in which the robot is able to recognize the previous location of the black disk in the three test environments. Groups of histograms indicate the performance for the corresponding location of the black disk (see Figure 12). White histograms indicate the performance obtained in simulation by moving the robot close to the light at the beginning of the second phase. Grey histograms indicate the performance obtained in simulation by moving the robot in a randomly selected location and orientation in the environment. Black histograms indicate the performance observed in hardware on the real environment by moving the robot in a randomly selected location and orientation. Left: data for the T-Maze environment. Centre: data for the simplified T-Maze environment. Right: data for the open arena environment with two light bulbs.

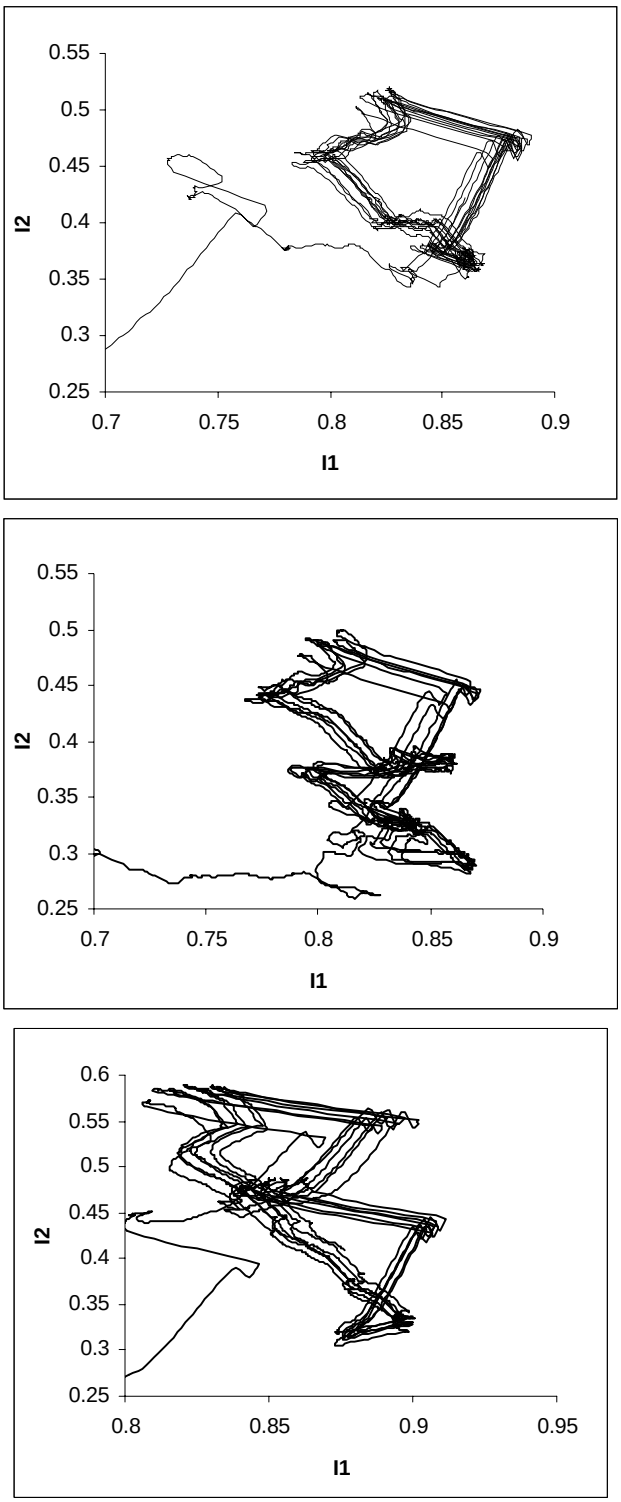


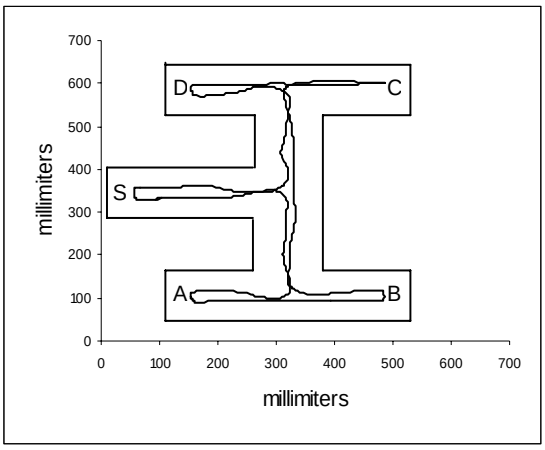
Figure 14**.** The activation state of the two internal neurons recorded while the robot moves in the environment for 330 s. Top: data for the T-Maze environment. Centre: data for the simplified T-Maze environment. Bottom: data for the open arena environment.

4.3. Robots' ability to discover elementary categorical features

As we have shown above, the categorical states generated by the robots are based on the combination and the integration of more simple categorical features. For example, the internal states which encode the four possible locations of the target areas (see Figure 6), are generated by exploiting fixed point attractors corresponding to selected elementary categorical features which ultimately correspond to specific sensory states (e.g. a sensory state encoding a light on the frontal side of the robot, or a sensory state encoding the presence of obstacles on the robot's left, frontal, and right sides, see Figure 8 and 9). Evolved robots, therefore, generate their internal state by integrating over time selected categorical features. These categorical features, in fact, depend on the control parameters which regulate how robots' sensory states affect robots internal states and on the control parameters which regulate how the robot moves in the environment which in turn co-determine the sensory states experienced by the robot. In this section we will focus on this aspect --- robots' ability to discover categorical features which can be integrated over tome so to generate effective spatial representation. In particular, by comparing the results of the basic experiment described above with the results of a new experiment in which robots are placed in double-T maze environment with the light turned off (i.e. in an environment which does not contain some of the elementary features exploited in the solution described above), we will show how evolved robots are able to select other elementary categorical features and to exploit them to generate an appropriate internal dynamics.

Figure 15 (right) shows the results obtained in the new experiment in which all the parameters were identical to that of the basic experiment but in which the light was turned off (i.e. in which an important landmark which was exploited in the previous experiments for self-localizing is missing). As can be seen, the evolved robot has a performance level which is similar to that of the basic experiment (with the exception of location D, which is recognized correctly only in about half of the cases).

By looking at the motor trajectory produced by the robot in the environment without the light, we can notice how the robot moves by closely approaching its left-side wall when it is situated close to location C (see Figure 15, left). This behaviour, which produces a sustained activation of the left infrared sensors of the robot in this part of the environment, is produced by the robot after a right turn preceded by a navigation trough a long corridor --- a situation which only occurs in this part of the environment for robots which produce a counter-clockwise wall-following behaviour. As shown in Figure 16, the perceptual event caused by this behaviour (i.e. the perception of a sustained level of activation on the left infrared sensors) corresponds to one of the fixed point attractors in the robot's internal dynamics which contribute to produce the periodic limit cycle dynamics shown in the Figure which in turn allow the robot to discriminate different environmental locations.



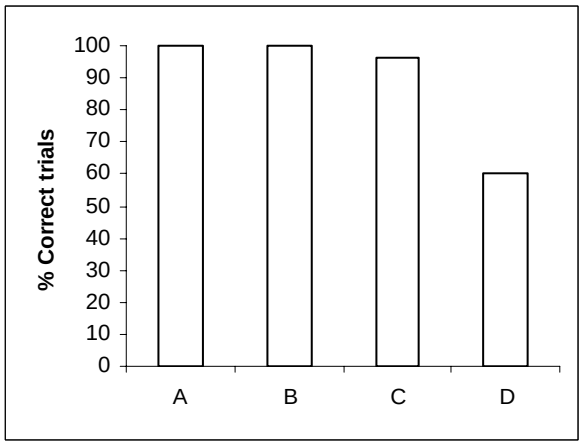


Figure 15**.** Result obtained in one of the best replications of the experiment conducted in a double-T maze environment with the light turned off. **Left**: Motor behaviour displayed by the best evolved robot of the best replication of the experiment. **Right**: Performance for different location of the black disk. Data obtained by moving the robot in position S after the target has been removed (see left Figure).

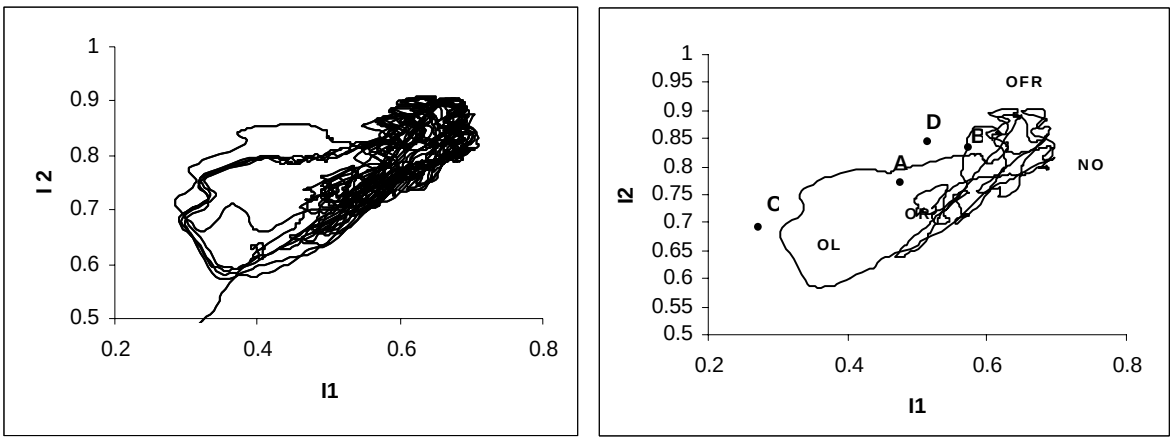


Figure 16. Experiment with the light turned off during the evolutionary and testing phase. **Left**: **T**he activation state of the internal neurons recorded while the robot moves in the environment for 330s. **Right**: schematization of the limit cycle shown in the left part of the figure. Labels A, B, C, D, and S indicate the internal states of the two neurons when the robot reaches the corresponding locations (see Figure 2). The OFR, OR, OL, and NO labels indicate the average fixed point attractors of five types of corresponding sensory states experienced by the robot when it detects: an obstacle on its right and frontal sides (OFR), the robot detects an obstacle on its right side (OR), the robot does not detect obstacles (NO), the robot detect an obstacle on its left side (OL).

Overall, the analysis reported above and the comparison with the strategy found in the experiment in which the light was turned on, indicate that the elementary categorical features that are used by the evolved robots to regulate their internal dynamics are selected by the robots during the adaptation process and are generated by the robot through the regulation of the robot's motor behaviour (i.e. are not only determined by the environment and by the robot sensory system but also by the robot's behaviour). Indeed, in the experiment reported in this section in which the environment provides fewer perceptual features with respect to the basic experiment, the evolved robot manages to create additional perceptual features through the exhibition of a specific motor behaviour. In other words, robots generate different elementary features and select the number and the type of features that can be appropriately used to solve the given problem.

8. Discussion

In this paper, we described how evolved robots provided with dynamical neural controller solve a problem that requires to identify their own relative location in the environment and to recognize previously visited locations.

The analysis of the obtained results indicates that evolved robots come up with simple and robust solutions that exploit limit cycle dynamics emerging from the coupling between the robot/environmental dynamics and the robot's internal dynamics.

The robot/environmental dynamics originates as the result of the effects that the actions performed by the robot produce on the sensory states which the robot will experience later on and as a result of the fact that the experienced sensory states co-determine the robots' successive actions. In the case of the

experiments described in this paper, the interaction between the robot and the environment, mediated by simple control rules which produce a form of wall following behaviour, leads to a sensory-motor flow that is quasi-periodic (over a time scale of about 50 s corresponding to a complete lap of the environment) and which is characterized by the alternation of few relatively stable type of sensory states lasting for different time duration (over a time scale ranging approximately between [0.5, 5] s).

The robot's internal dynamics originates as a result of the fact that the current state of the internal neurons affects their future state. In the case of the experiments described in this paper, the robot's internal dynamics is realized through leaky integrator neurons, i.e. neurons which hold a certain amount of their activation from the previous time step. The output of these neurons tend to approximate different fixed point attractors, corresponding to different values of the net input contribution received from other connected neurons, with a tuneable time scale which is determined by the time constant parameter of the neurons and which vary in the range [0, 100] s.

The interaction between the robot/environmental dynamics and the robot's internal dynamics is due to the fact that the former dynamics determines the way in which sensory states vary in time which, in turn, affects the latter dynamics by determining the way in which the corresponding fixed point attractors in robot's internal state space vary in time. The possibility of generating a coupled robot/environmental and robot's internal dynamics is ensured by the fact that the parameters that regulate the characteristics of the two dynamics are encoded into free parameters that are co-adapted and co-shaped during the evolutionary process. These parameters affect in particular the time scale at which the robot/environmental and the robot's internal dynamics occur, the type of sensory states that are experienced, and the relation between the experienced sensory states and the corresponding fixed point attractors in the robot's internal dynamics.

The analysis of the obtained results indicates that the coupling between the two dynamical processes leads to limit cycles in the state space of robot's internal neurons which are characterized by expressiveness (i.e. by the fact that different internal states correspond to different spatial locations of the robot in the environment) and periodicity (i.e. by the fact the same spatial location corresponds to the same internal state independently from the characteristics of previously experienced sensory states). This limit cycle dynamics is the result of a transient dynamical process that originates from the fact that the robot's internal state approximates the current fixed point attractors by never fully reaching it since fixed point attractors and their corresponding sensory states last for limited time durations and since the state of the internal neurons vary slowly.

The strategy discovered by evolved robots is robust and remarkably general, in the sense that it allows the robot to generalize their skill in a wide variety of circumstances that were not experienced by the robot during the evolutionary phase.

The results described in this paper also provide a contribution for what concern the question of how robots can represent spatial information (e.g. the position of selected locations of the environment) and how they can self-localize on the basis of the information collected while they move in the environment. Indeed, it is worth noting that the characteristics and the mechanisms that allow our evolved robots to display these skills differ qualitatively from other models described in the literature. Such specificity consists in the fact that the way in which selected spatial locations and the overall spatial organization of the environment are represented in robots' internal states is implicit rather than explicit as in other models (e.g. metric or topological maps approaches in which the geometrical features of the environment or the spatial relations between relevant locations of the environment are represented in 2D maps or in topological graphs, Meyer & Filliat, 2003). In explicit representation systems, such as those reviewed in the paper referred above, different spatial locations are encoded by different elements of the robots' control systems (e.g. nodes and connections between nodes) which are activated or not depending on whether the robot is located into the corresponding physical location. In implicit representations systems instead, such as those developed by our evolved robots, different

spatial locations are encoded by the same elements of the robots' control system through different activation states. This implies that, while in the former case the elements that potentially represent different spatial locations exist independently from the actual position of the robot and are activated when the robot reach the corresponding location, in the latter case the representations are generated while the robot is situated in the environment.

One important implication of the difference between implicit and explicit representation forms is that, as we have shown in this paper, implicit representations can allow robots to generalize their ability in new environments immediately without the need to acquire an explicit representation of the new environment. In the case of implicit representation forms, in fact, the new representation corresponding to the new environment is generated immediately as soon as the robot is situated in the new environment without the need to change the free parameters of the system.

Acknowledgment

This work was supported by the ECAgents project, a project funded by the Future and Emerging Technologies programme (IST-FET) of the European Community, under grant 001940. 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

Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, *11*, 209–243.

Beer, R.D. (1995a). A dynamical systems perspective on agent-environment interaction. Artificial Intelligence 72:173-215.

Beer, R.D. (1995b). On the dynamics of small continuous-time recurrent neural networks. Adaptive Behavior 3(4):471-511.

Brooks R.A. (1991) Intelligence without reason. In J. Mylopoulos & R. Reiter (Eds.), *Proceedings of 12th International Joint Conference on Artificial Intelligence*. San Mateo, CA: Morgan Kaufmann.

Chiel, H.J. and Beer, R.D. (1997). The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. Trends in Neurosciences 20:553-557.

Clark A. (1997) *Being There: Putting Brain, Body and World Together Again.* Cambridge, MA: MIT Press.

Clark, A. & Thornton, C. (1997) Trading spaces: Computation, representation, and the limits of uniformed learning. Behavioral and Brain Sciences 20:57-90

Floreano D. & Mondada F. (1996) Evolution of homing navigation in a real mobile robot. *IEEE Transactions on Systems, Man, and Cybernetics--Part B: Cybernetics*, 26(3):396-407.

Morimoto, G. and Ikegami, T. (2004). Sensory-motor coupling and dynamical categorization'. In J.Pollack et al. (Eds.) Proceedings of ALIFE IX. Cambridge, MA: MIT press.

Meyer J-A., Filliat D. (2003). Map-based navigation in mobile robots: II. A Review of map-learning and path planning strategies. Cognitive Systems Research, 4: 283-317.

Mondada R., Franzi E. & Ienne P. (1993). Mobile robot miniaturization: A tool for investigation in control algorithms. In T.Y. Yoshikawa & F. Miyazaki (Eds.), *Proceedings of the Third International Symposium on Experimental Robots*. Berlin, Springer-Verlag.

Nolfi S. & Floreano D. (2000). Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines. Cambridge, MA: MIT Press/Bradford Books.

Nolfi S. & Marocco D. (2001). Evolving robots able to integrate sensory-motor information over time. Theory in Biosciences, 120:287-310.

Nolfi S. (2002). Evolving robots able to self-localize in the environment: The importance of viewing cognition as the result of processes occurring at different time scales. Connection Science (14) 3:231-244.

Nolfi S. (2005a). Behaviour as a complex adaptive system: On the role of self-organization in the development of individual and collective behaviour. *ComplexUs*, 3-4: 195-203.

Nolfi S. (2005b). Categories formation in self-organizing embodied agents. In H. Cohen & C. Lefebvre (Eds), *Handbook of Categorization in Cognitive Science*. Elsevier.

Nolfi S., and Tani J. (1999). Extracting regularities in space and time through a cascade of prediction networks: The case of a mobile robot navigating in a structured environment, *Connection Science*, (11) 2:129-152.

Pfeifer R. and Scheier C. (1999) *Understanding Intelligence*. Cambridge, MA: MIT Press

Slocum, A. C., Downey, D. C., & Beer, R. D. (2000). Further experiments in the evolution of minimally cognitive behavior: From perceiving affordances to selective attention. In J. Meyer, A. Berthoz, D. Floreano, H. Roitblat, & S. Wilson (Eds.), *From Animals to Animats 6: Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior*. Cambridge, MA: MIT Press.

Tani J., and Fukumura N. (1997). Self-organizing internal representation in learning of navigation: A physical experiment by the mobile robot Yamabico. *Neural Networks*, 10 (1): 153-159.

Tani J., and Nolfi S. (1999). Learning to perceive the world as articulated: An approach for hierarchical learning in sensory-motor systems. *Neural Networks*, 12:1131-1141

Trullier O., Wiener S.I., Berthoz A., Meyer J-A. (1997). Biologically based artificial navigation systems: Review and prospects. *Progress in Neurobiology* 51:483–544.

Varela F.J., Thompson E. and Rosch E. (1991) *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Vickerstaff, R. J., and Di Paolo, E. A., (2005a). An evolved agent performing efficient path integration based homing and search. Proceedings ECAL 2005, Springer.

Vickerstaff, R. J., and Di Paolo, E. A., (2005b). Evolving neural models of path integration. *Journal of Experimental Biology*, 208: 3349-3366.

Wolpert, D., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. Neural Networks, 11. 1317-1329.

About the authors

Onofrio Gigliotta is a Ph.D student at the University of Palermo and research assistant at Institute of Cognitive Sciences and Technologies of the Italian National Research Council (CNR-ISTC). His

research interests include evolutionary robotics, spatial navigation, and social simulations within the general framework of a dynamical approach to cognition.

Stefano Nolfi is director of the Laboratory of Autonomous Robots and Artificial Life within the Institute of Cognitive Sciences and Technologies of the Italian National Research Council (CNR-ISTC). He published more than 100 peer-reviewed articles and a book on Evolutionary Robotics. His research activities focus on Evolutionary Robotics, Collective Robotics, Adaptive Behavior, Complex Systems, Embodied Cognitive Sciences.