Evolving Robots Able To Integrate Sensory-Motor Information Over Time

Stefano Nolfi[1] Davide Marocco[2]

[1]Division of Neural Systems and Artificial Life, Institute of Psychology, National Research Council (CNR), Viale Marx, 15, 00137, Rome, Italy
nolfi@ip.rm.cnr.it
[2] University of Calabria, "Centro Interdipartimentale della Comunicazione"
87036 Arcavacata di Rende, Cosenza, Italy
davidem@ip.rm.cnr.it

Abstract

We will discuss in which conditions we can expect the emergence of agents able to integrate sensory-motor information over time and later use this information to modulate their behavior accordingly. In doing so we will illustrate the problems that these agents should be able to solve and the processes that might lead to a transition from simple agents that only rely on sensory information or on their internal dynamic to agents that are also able to integrate information over time. The analysis of evolved individuals revealed that: (1) individuals able to integrate information over time rely on mixed strategy in which basic sensory-motor mechanisms are complemented and enhanced with additional internal mechanisms; (2) evolved individuals tend to rely on partial, action-oriented, and action-mediated representations of the external environment.

1 Introduction

A new research paradigm, that has been called *Embodied Cognitive Science*, has recently challenged the traditional view according to which intelligence is an abstract process that can be studied without taking into consideration the physical aspects of natural systems. In this new paradigm, research tends to stress the importance of 'situatedness' (i.e., the importance of studying systems that are situated in an external environment) and embodiment (i.e., the importance of study systems that have bodies, receive input from their sensors and produce motor actions as output) (Varela, Thomson, and Rosch, 1991; Brooks, 1991; Hendriks-Jansen, 1996; Clark, 1997; Pfeifer and Scheier, 1999). An attractive way to study systems that are embodied and situated is *evolutionary robotics,* i.e. the attempt to develop robots through a self-organized process based on artificial evolution (Nolfi and Floreano, 2000). This approach stresses the importance of study system that have a real body, are situated in a physical environment, and develop their own skills in close interaction with the environment without human intervention.

Current research in evolutionary robotics often involves agents that are mostly reactive or that are mostly driven by their internal dynamics. The former are agents in which motor actions are based, mostly, on the current state of the sensors and in which internal states, if any, play a comparatively limited role. The latter are systems that are primarily driven by their internal dynamic and in which sensory information, if any, play a comparatively limited role. An example of the former category is the garbage collecting robot developed by Nolfi (1997a, 1997b). This robot is able to display a rather long sequence of non-trivial behaviors

(i.e., explore the environment, avoiding walls; recognize a target object and move to an angle and distance from where the object can be grasped; pick up the target object; move toward the walls while avoiding other target objects; recognize a wall and move at an angle and distance from where the object can be safely dropped outside the arena; release the object). In order to achieve its goal however, the robot does not need to integrate information over time. Indeed a simple reactive controller that always react in the same way to the same sensory states turns out to be sufficient. An example of the latter category is the walking animat developed by Beer and Gallagher (1992). The agent is able to display an ability to walk based on its own internal dynamic (i.e., on the ability to generate the right ritmic movements of the legs through time) without the need to also take into account sensory information (i.e., on the basis of a neural network controller that does not had access to any sensory information). The author also showed how systems that also rely on sensory information can produce a more effective walking behavior and display interesting properties. We will come back on this point later on. For the moment what it important to notice is that this agent and several other walking machines that have been developed later on (Lewis et al. 1992; Jacoby, 1998; Gomi and Ide, 1998) mostly or fully rely on their internal dynamic (i.e., the ability to integrate sensory-motor information over time play a limited role).

In this paper we will discuss in which conditions we can expect the emergence of systems able to integrate sensory-motor information over time and later use this information to modulate their behavior accordingly (for a similar view see [Beer, 1995, 1996]). In doing so we will discuss the problems that these systems should be able to solve and the processes that might lead to a transition from simple agents that only rely on sensory information or on their internal dynamic to system that are also able to integrate information over time.

In section 2 we will discuss what we mean by agents that are able to integrate sensory-motor information over time and the relation of such a notion with the notion of internal representation. In section 3 we will discuss the possibility to identify classes of problems that require an ability to integrate information over time. In section 4 we will describe which neural substrates might be suitable to allow the emergence of such systems. In section 5 we will try to identify the challenges that evolving individuals should face to develop an ability to integrate information over time. In section 6 and 7 we will present two experiments in which evolving robots develop an ability integrate sensory-motor information over time and later use this information in order to modulate their behavior accordingly. In section 8 we will discuss the implications of the obtained results.

2. On the notion of internal representation

One possible way to characterize agents that are able to integrate information over time is to resort on the notion of *internal representations.* This term is used to indicate internal states that encode information about the external environment that are not directly available in the current sensory state and that have the function of carrying specific types of information (Clark, 1997). Agents that exploit internal representations as well as information directly available from their sensors and that are able to extract their internal representations autonomously by interacting with the environment are, by definition, agents that are able to integrate sensory-motor information over time.

Unfortunately, there is no a formal way that can be used to identify whether internal states are internal representation. This can be explained by considering that:

(a) There is no clear way to identify whether an internal state has a representational status or whether it simply display an accidental correlation with some feature of the external environment (Beer, 1995; Clark 1997).

(b) Natural organisms and artificial organisms that are left free to self-organize tend to have internal states that are local and action-oriented rather than objective and action independent. Identifying what is the relation between such states and the external environment is rather difficult especially from the perspective of an external observer. The problem is further complicated by the fact that, as we will show in the next sections, internal states tend to be action-mediated (i.e. they implicitly and indirectly encode information about the external environment and/or the agent/environment relation by encoding information on what the agent did in its previous interaction with the environment).
(c) The interaction between an individual and the external environment often can be properly described as a process of continuous reciprocal causation (Clark, 1997). In such cases, describing the individual and the environment in isolation and trying to identify what is the relation between the internal states of the agent and the features of the external environment might be not feasible.

To summarize the notion of internal representation can be more properly characterized as a description that is in the eye of the observer rather than as a formal property of an agent. For this reason we will resort on the more general notion of internal state. By *internal state* we mean a state (e.g. the activation state of an internal neuron of the control system of a robot) that might be affected by the previous sensory-motor states experienced by the robot and that co-determine, together with the current sensory states, the robot motor actions. By mediating between perception and actions, internal states might allow agents to produce behavior that are decoupled from the immediate circumstances while still remaining sensitive to them (Slocum et al., 2000). We will use *reactive robot* to indicate a robot that does not have any internal state and in which the current motor action is only function of the current sensory state. We will use *robot that only rely on its internal dynamic* to indicate a robot in which sensory information coming from the external environment is missing or do not play any role in determining the robot motor actions.

Internal states might consists of different entities. For example, in the case of a neural controller, they might consists in the activation state of some neurons and/or in the strength of the synaptic weights. It should be noted that there is not a one to one correspondence between the architecture of the controller and the type of strategy adopted by evolving individuals. For instance, although an individual provided with a recurrent neural network controller might potentially develop an ability to integrate information over time, it might also rely on a simple reactive strategy.

3 Identifying the class of problems that require to integrate information over time

Every one that had the chance to play with artificial evolution knows that it will find the easiest solutions first. From our point of view this means that, given that agents that rely on the sensory states only or on their internal dynamic only are simpler than agents that are also able to integrate information over time, we should always expect that the agents that emerge in the first phase of the evolutionary process very unlikely will display an ability to integrate information over time. Whether individuals able to do so will emerge in the successive phase or not depend on: (a) how much space there is for improvement, i.e. how strong it is the selective pressure and, (b) the evolvability of the agents themselves, i.e. the probability that the application of the genetic operators might lead to offspring with better performance. We will discuss (a) in this section and (b) in the next.

In order to study how agents able to integrate information over time can emerge we have to identify the problems that cannot be solved by agents that rely on reactive behavior only or on their internal dynamics only (see also Nolfi [in press], Nolfi and Marocco [in press]). At a first sight, one might think that many tasks such as landmark-based navigation and maze negotiation, require agents able to integrate information over time. As we will demonstrate however, identifying problems that can only be solved by agents able to integrate information over time would require a complete knowledge of all the possible ways in which a given problem can be solved and this is not feasible even in relatively simple cases.

One possible way to identify problems that cannot be solved without integrating information over time is to resort on the notion of *sensory aliasing*[1]. Sensory aliasing refers to the situation wherein, from the agent point of view, two or more agent/environmental states correspond to the same sensory pattern but require different motor responses. One way to solve this problem in fact, is to integrate the sensory pattern affected by the aliasing problem with a trace of the previously experienced sensory patterns. As claimed above however, to identify if the ability to solve a problem involves sensory aliasing one should know all possible ways to solve that problem.

Consider a simulated agent which lives in a circular stripe divided into 40 cells (20 on the left and 20 on the right side). Cells are labelled by numbers in the range 0-19 in different random orders (see Fig. 1). At each time step the agent occupies one single cell and perceives the number corresponding to that cell. The agent can react to the current sensory state by moving one cell clockwise or counter-clockwise. Starting at any position on the circular stripe, the goal of the agent is to reach and/or remain in the left part of the environment.

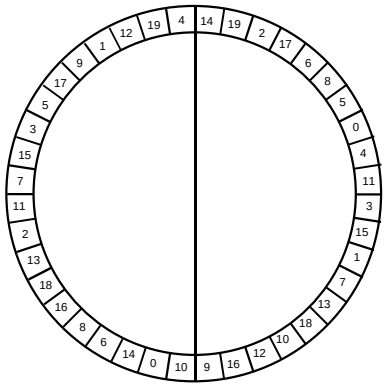


Fig. 1. The environment. The numbers represents the sensory state experienced by the agent in each cell. Each different sensory state is present one time in the left and one time in the right part of the environment, but at different random location.

Apparently, in this case all sensory states are affected by the aliasing problem (whatever the sensory state is, in fact, the agent has 50% of probability to be in the left or in the right part of the environment) and the only possible solution is to integrate information over time. For example we might imagine an agent that turn clockwise by keeping track of the previous experienced sensory state and stop as soon as the current and previous sensory states are 0 and 10 respectively.

By evolving a population of agents for their ability to solve this task[2] however, one can easily see that there are several other solutions that are not affected by the aliasing problem

[1] The notion of sensory aliasing is related to the notion of perceptual aliasing introduced by Whitehead and Ballard (1991). Perceptual aliasing was introduced to indicate an internal state corresponding to different external states that requires different motor answers in different circumstances.

[2] Evolving individuals were allowed to “live” for 100 epochs with each epoch consisting of 200 actions. Each

and therefore do not require to integrate information over time (Nolfi and Parisi, 1999). One of these solutions is shown in Fig. 2. Individuals react to the same sensory state always in the same manner. Despite of that, the way in which they react to the different sensory states allow them to produce behavioral attractors in the left but not in the right part of the environment. Attractors consists of two adjacent cells to which the agent react clockwise and counterclockwise (following the clockwise direction, the robot should respond clockwise to the first cell and counterclockwise to the second cell, see points indicated with an “a” in Fig. 2). When the agent encounters an attractor point, it starts to move back and forth while remaining there. This allows individuals to abandon the right side of the environment by consistently move clockwise or counterclockwise at to remain in the left side of the environment by oscillating on one of the attractors. For an example in which the same type of strategy emerges by evolving a Khepera robot (Mondada, Franzi and Ienne, 1993) for the ability to discriminate between object with different shapes see Nolfi and Parisi (1999).

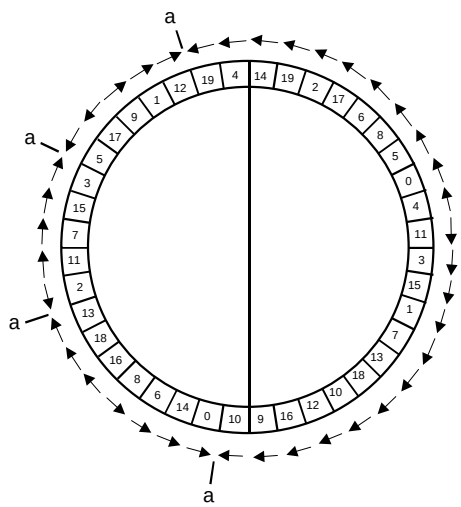


Fig. 2. A typical strategy obtained with artificial evolution. Arrows indicate the motor reaction of an evolved agent. Points indicated with an “a” are attractors in the behavior of the agent (i.e., points in which the agent starts to oscillate by moving clockwise and counterclockwise).

The impossibility to predict all the possible solutions of a given problem prevent the possibility to resort on the notion of sensory aliasing to identify the tasks that are suited for the evolution of agents able to integrate information over time.

One might argue that simple reactive solutions, such as that described above, exploit surface regularities of the particular environment in which they have been selected but will not generalize in different environmental circumstances. In other words, one might be tempted to conclude that these are tricky solutions that are only valid in very special circumstances.

To show that this hypothesis does not necessarily hold, consider the case of a Khepera robot placed in a randomly selected location within a rectangular environment that should

epoch individual were initially placed in a randomly selected cell of the environment. Individuals' fitness was increased each time individuals end up in the left part of the environment after 200 cycles. Individual's controller consisted of a neural network with 20 sensory neurons (locally encoding the corresponding sensory state) and a motor neuron (encoding one of the two possible actions: move clockwise or counterclockwise) Connection weights were represented in the genotype with 20 bits (1 bit for each connection). The population size was 100. The best 20 individuals of each generation were allowed to reproduce by generating 5 copies of their genotype with 2% of their bits replaced with a new randomly selected value. The experiment was replicated 10 times.

navigate toward the north-west or the south-east corners of the environment (a similar task was studied by Gallistel [1990] on rats). The length of short walls and the proportion between long and short walls was randomly varied in each epoch (short walls varied between 15 and 45 cm, long walls varied between 16 and 90) (see Fig. 3).

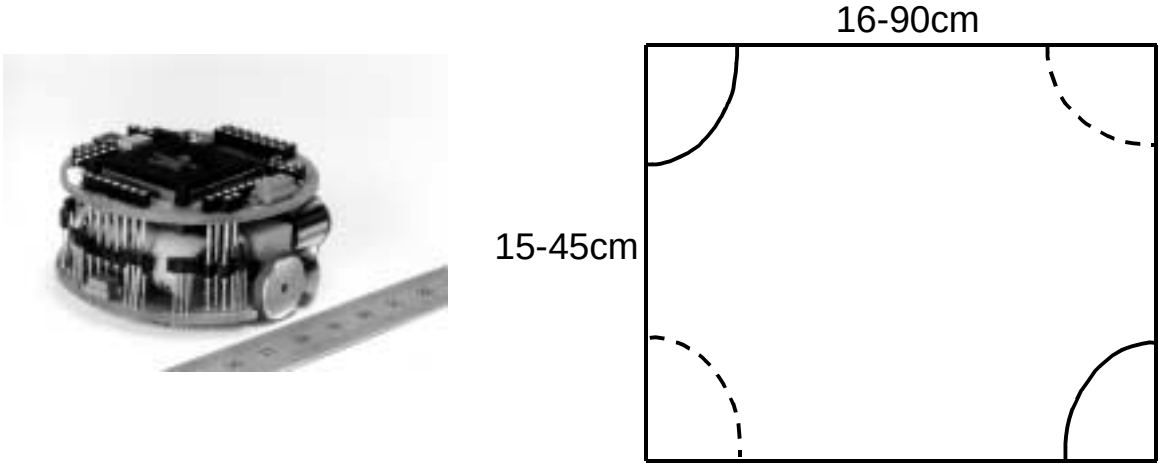


Fig. 3. Left: The Khepera robot. **Right:** The environment. The lines represent the walls. Full and dashed arcs represent the right and wrong areas respectively.

The robot is provided with 8 infrared sensors that can detect the walls at a maximum distance of approximately 4 cm and two motors controlling the two wheels (see Fig.3, left). Apparently, the only way to solve this task is to discriminate between long and short walls and to navigate accordingly. For example the robot could follow a long wall on its own right side until a corner (or other similar strategies like following a short wall on its own left side until a corner). Given that sensors can only provide information about the local portion of the environment surrounding the robot, the ability to discriminate between long and short walls seem to necessarily require an ability to integrate information over time.

By running an evolutionary experiment[3] however one can easily see that even robots controlled by a pure reactive controller with 8 sensory and 2 motor neurons (see Fig. 4, left) can solve this problem to a good extent (up to 85% of correct navigations in the case of the best replications of the experiment). In other words, simple reactive strategies might be effective (although not optimal) even when agents are asked to accomplish a task in rather different environmental conditions.

Fig. 4 shows a typical strategy displayed by evolved individuals. The robot explores the environment by avoiding walls and by moving forward in straight lines until it finds one of the corners (in the case shown in the Fig. the robot starts from the right side of the

[3] Robots were provided with a purely reactive neural controller consisting of a fully connected perceptron with 8 sensory neurons (encoding the state of the 8 infrared sensors of the robot) and 2 motor neurons (encoding the speed of the two motors connected to the two corresponding wheels). Individuals' fitness was increased or decreased of 1 point each time individuals ended their lifetime in one of the two right or wrong corners respectively. The genotype encoded the connection weights and biases of the neural controller. Each weight was represented in the genotype by a 8-bit string and normalized between –10 and +10. Population size was 100. The best 20 individuals of each generation were allowed to reproduce by generating 5 copies of their genotype with 4% of their bits replaced with a new randomly selected value. Lifetime consisted of 24 epochs each consisting of 500 lifecycles. At the beginning of each epoch, the length of short walls and the proportion between long and short walls was randomly set (the length of short walls was selected randomly between 15 and 45 cm and the length of long walls was obtained by multiplying the length of short wall by a number selected randomly between 1.1 and 2.0). Moreover, the robot was placed in a new randomly selected position and orientation. The evolutionary process was conducted in simulation by using the Evorobot software (Nolfi, 2000).

environment and encounters the bottom-left corner first). When it is facing a corner, the robot moves left and right and back and forth for a while and then leaves the corner with an angle of about 45° with respect to the two walls forming the corner. Finally, when it encounters a wall with an angle of about 45° on its left or right side it turns counter-clockwise until it has the wall on its right side and then follows the wall until the corner.

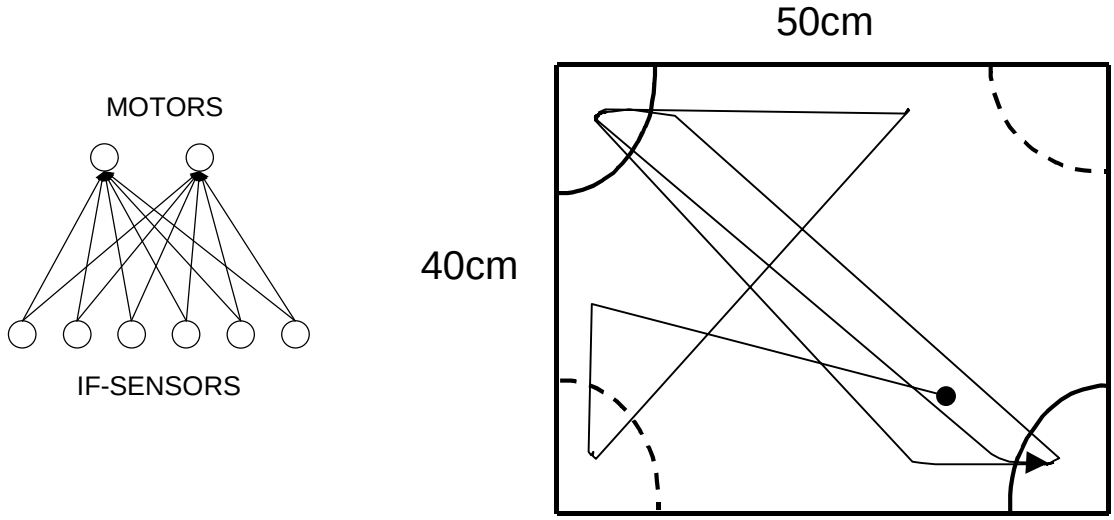


Fig. 4. Left: The architecture of the neural controller of the robot. **Right:** A typical strategy displayed by evolved individuals. The full circle and the arrow represents the initial and final position of the robot respectively during an epoch. The line represents the trajectory of the robot. The full and dashed arcs represent the areas in which the robot is rewarded or punished respectively.

This strategy ensures that, after the first corner (that might be any corner given that the initial position and orientation of the robot is randomly chosen), the robot will always move toward one of the two right corners. Notice that this strategy works for any rectangular environment independently of the relative dimensions of the long versus short walls and of the absolute dimension of the environment. Indeed leaving corners with an angle of 45° assures that the robot will always encounter one of the two longer walls next. Interestingly this strategy allows the robot to "measure" the relative length of the walls at once. In other words, the robot should not measure two adjacent walls and compare their length in order to ascertain which is the longer and the shorter wall.

We will come back on the limits of this strategy that prevent the achievement of optimal performance in section 5 where we will discuss how this strategy can be improved by integrating information over time.

To summarize we might conclude that there are no clear ways to identify the cases that necessarily need agents able to integrate information over time and later use such information to achieve their goals. This can be explained by considering that the type of internal processes needed depend on the strategy adopted to solve the task and, from an external observer, it is impossible to identify all the possible ways in which a task can be solved.

4 Identifying a neural substrate suitable for the evolution of agents able to integrate information over time.

The second question we raised in section 3 is how to maximize the evolvability. The probability to obtain individuals able to solve tasks that require an ability to integrate information over time, in fact, also depends on the probability that the application of the

genetic operators can produce better individuals and this in turn depends on the parameters that are subjected to the evolutionary process. In this section we will restrict our analysis to the agents' controllers although the type of sensors and actuators and the body structure might obviously play an important role.

Most of the experiments in evolutionary robotics rely on neural controllers. In many cases feed-forward neural networks are used. These networks are effective in producing reactive behavior but cannot deal with time (i.e., they always react in the same way to the same sensory state and therefore cannot integrate information over time).

In other cases recurrent neural networks (Jordan, 1989; Elman, 1990) have been used. By including recurrent connections these networks can produce responses that take into account both the current and the previous experienced sensory states. These networks have been successfully applied to a navigation task in which a robot had to periodically return to a charging area (Floreano and Mondada, 1996a). However, the general effectiveness of these type of network in dealing with real-time events and the ability to scale up is unclear.

Other attempts have been conducted by using continuous time recurrent neural network (Beer and Gallagher, 1992). By relying on differential equations instead of being updated at fixed time steps these networks can produce continuous dynamics. These networks have been successfully applied to variety of tasks (such as legged locomotion [Gallagher et al. 1998] and visually guided navigation [Harvey et al. 1994]). However the extent to which they can be applied to tasks that have sequential components and their ability to scale up is unclear.

Other attempts have been conducted by using synaptic plasticity. In some cases the synaptic weights were updated by using reinforcement learning (Ackley and Litman, 1991) or back-propagation (Nolfi and Parisi, 1993) on the basis of self-generated teaching signals. In other cases synaptic weighs were updated on the basis of genetically encoded hebbian rules (Floreano and Mondada, 1996b). In general terms, integration of information over time can be accomplished both by modifying the synaptic weights (through some form of plasticity) and by means of recurrent connections. In both cases in fact, the way in which individuals react to the current sensory state might be affected by the previous experienced sensory states. Differ methods however might have different characteristics. For instance, the former approaches, by relying on gradient descent technique, tend to produce small and long term effects on the robot behaviors whereas the latter approach, based on hebbian learning, might produce significant effects in the short term (see Floreano and Mondada, 1996b).

Still other possibilities include neural network models that include modulatory mechanisms such as the GAS network proposed by Husband et al. (1999).

Unfortunately, a systematic comparison of the advantages and disadvantages of each neural network type has not been carried out jet.

5 Challenges

Apart from the fact that simple solutions that only rely on sensory-motor coordination or on the agent's internal dynamic might constitute a local minimum from an evolutionary perspective, the difficulty of evolving individuals able to integrate information over time can be explained by considering the complexity of the problem itself. In this section we will try to identify which are the most serious challenges.

A first problem arises from *the need to develop two abilities at the same time*. Given that during the first phase of the evolutionary process individuals tend to select simple strategies that rely on the current sensory states only or on the internal dynamic of the agent only, developing an ability to integrate information over time require to: (a) extract useful internal states from the sensory-motor flow and (b) modulate behavior appropriately on the basis of

the extracted internal states. Internal states in fact do not provide any selective advantage by themselves. Similarly, an ability to appropriately rely on internal states without an ability to reliably extract such states does not provide any adaptive advantage. From an evolutionary perspective this implies that the two abilities should be developed at the same time.

A second hard challenge arises from *the need to couple the internal dynamic with the sensory-motor flow*. Behaving agents receive a continuous flow of sensory-motor states. From this huge amount of information they should be able to extract a limited number of states that capture useful features of the agent/environment interaction and that can be used to later affect their behavior. Beside of that, agents should be able to continuously update their internal states by taking into account changes occurring in the agent/environment relation and/or in the environment itself. In other words, internal states should change accordingly to a given dynamic that should be coupled with the dynamic of the sensory-motor flow.

A successful example of coupling between the agent's internal dynamic and the sensory-motor flow has been described by Beer (1995). In his experiments on six-legged walking animats, the author showed how evolved continuous time recurrent neural networks were able to produce an effective walking behavior on the basis of their internal dynamics only (i.e., without any sensor information). Networks also provided with sensory information however displayed better walking and spontaneously exhibited an ability to modulate the walking behavior on the basis of the sensory-motor flow, i.e. evolved networks displayed an ability to speed up or slow down the walking behavior when the rate of change of the sensors was artificially increased or decreased. It should be noted, however, that in this case sensory information and motor responses varies at the same time scale and sensory information is always available. In other cases, relevant sensory information might be available only time to time and the dynamics of the relevant sensory-motor flow and of the motor effectors might occur at different time scales. These factors might significantly complicate the coupling between the two dynamical processes.

Finally, a third challenge arises from the fact that the exploitation of previously extracted internal states might increase the *unreliability* of individuals. Sensory and motor states are not objective measures and are affected by a large number of unpredictable variables (noise, ambient light, battery level etc.). Internal states however, in addition to that, also depend on the behavior previously displayed by the agent (i.e., individuals might come up with the wrong internal states). Relying on internal states that have been previously extracted therefore might introduce stochastic elements that are not present in agents that are only driven by sensory-motor states or by their internal dynamics (a similar claim has been made by Mayley [1997] with reference to plastic individuals).

Each of these issues might represent a rather serious challenge from an evolutionary point of view. Fortunately, as we will see in the next section, simple cases exist in which only some of these issues should be addressed. These cases might represent important stepping stones that might allow the evolutionary process to progressively move from simple agents that only rely on sensory-motor coordination or on their internal dynamics to more complex agents that are also able to integrate information over time.

6 Solving the landmark navigation task

Let us go back to the landmark navigation problem that we described in section 3. The simple solution that only rely on sensory-motor coordination has a limit that explains why evolved individuals fail in some cases although they are able to successfully navigate toward one of the two right corners in about 75% of the trials, on the average (see Fig. 5, thin lines), and up to the 85% of the trials in the best replications (result not shown).

The problem is that these individuals are unable to stop on the right corners (Fig. 4). They act so to avoid the wrong corners but once they reach one of the two right corners they keep travelling back and forth between the two right corners. In doing so they might fail to satisfy the fitness criterion, given that individuals gather fitness only if they end their lifetime in one of the two right corners. Actually these individuals display two behavioural attractors at different time scales. At a short time scale, individuals oscillate in front of a corner in order to remain there for a while and then leave. At a longer time scale, they keep oscillating between the two opposite corners. These oscillatory behaviours can be explained by considering that the robot, by being provided with a simple reactive neural controller, cannot "know" if the actual corner is one of the two right or one of the two wrong corners. By quickly travelling toward the right corners and by spending some time in each corner individual maximise their chance to end their lifetime in one of the two corners.

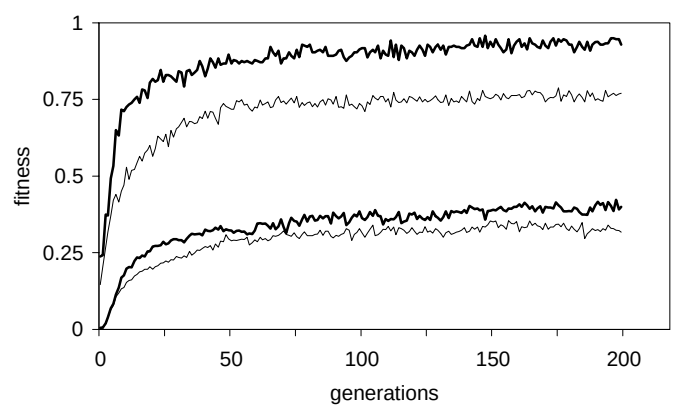
At this point we might try to ask the following question: in order to fully solve this problem, should evolving individuals select a significant different strategy or might they complement this basic strategy with additional mechanisms?

To answer this question we replicated the experiment by providing the robots with neural network controllers with two additional internal neurons (receiving connection from the sensory neurons and the internal neurons themselves and projecting connections to the two motor neurons)[4]. The activation state of sensory and internal neurons was updated accordingly to the following equation (motor neurons were updated according to the standard logistic function):

$$A_j = t_j + \sum w_{ij} O_i \qquad O_j = \tau_j O_j^{(t-1)} + (1-\tau_j)(1+e^{-A_j})^{-1} \qquad 0 \le \tau_j \le 1 \qquad \textbf{(1)}$$

With *Aj* being the activity of the *j*th neuron (or the state of the corresponding sensor in the case of sensory neurons), *tj* 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. *Oj* is the output of the *j*th neuron, τ*j* the time constant of the *j*th neuron.

Individuals provided with such additional internal layer reach close to optimal performance on the average (see Fig. 5), optimal performance in the case of the best replications (results not shown), and significantly outperform individuals provided with simple reactive neural controllers (see Fig. 5).



[4] The genotype encoded the connection weights and biases of the neural controller and the time constant parameters of sensory and motor neurons. Each parameter was represented in the genotype by a 8-bit string. Weights and biases were normalized between −10 and +10. Time constants were normalized between 0.0 and 1.0. All other parameters were the same of the experiment described above.

Fig. 5. Average fitness of the population and fitness of the best individuals along 200 generations for individuals with or without the additional layer of neurons (thick and thin lines respectively). Average result of 10 replications of the experiment.

Interestingly the exploitation of the internal states does not imply a profound re-organization of the strategy itself. Fig. 6 shows a typical behavior displayed by evolving individuals in which you can easily recognize the same strategy that we described in section 3 (i.e., individuals leave corners with an angle of about 45°, encounter one of the two long walls, and then follow the wall until the corner by keeping the wall on their right side). The only difference is that these individuals are able to stop on the right corners. This is accomplished by exploiting the internal dynamics of the neural controller and by using internal states to modulate the reactive behavior of the robot. In particular, in most of the evolved individuals one can observe an internal unit whose activation increases up to the maximum during a certain amount of time and then remains stable (in the case of the individual displayed in Fig. 6 this is the case of the unit h1). This unit modulates the behavior of the robot by letting it leave corners during the first phase in which the robot negotiates with the environment and allowing it to remain on corners after such initial phase.

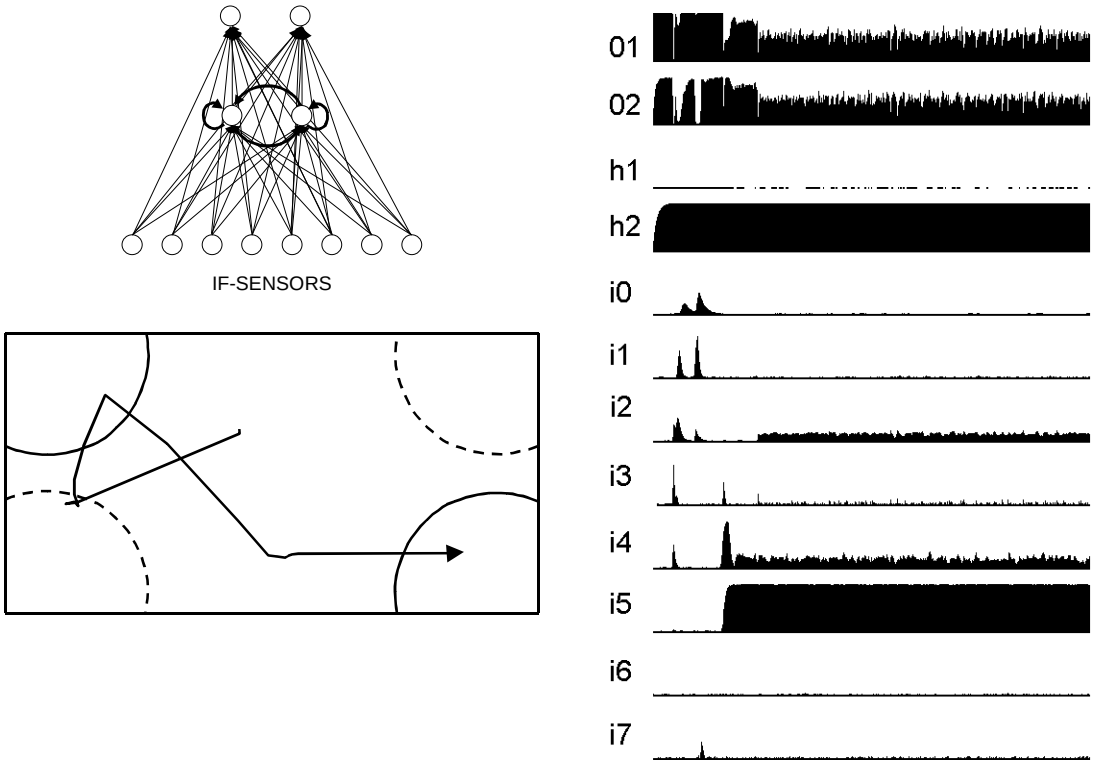


Fig. 6. Top, **left:** the architecture of the neural controller. **Bottom**, **left:** typical behavior displayed by an evolved robot provided with the additional layer of internal neurons. **Right**: state of the two motor neurons (o0 and o1), two internal neurons (h0 and h1) and 8 sensory neurons (from i0 to i7) throughout time of the individual displayed on the left side of the figure.

The analysis of obtained results therefore suggests that more complex individuals that are able to fully solve the problem mostly rely on the same sensory-motor strategy discovered by reactive individuals. Such a strategy is modulated by an extremely simple additional internal mechanism that just keep track of how much time the robot has spent negotiating with the environment.

7 Visually-guided robots able to discriminate between landmarks of different size.

A similar picture emerge from another experiment in which robots were asked to discriminate between visual landmarks by navigating toward the larger and not the smaller visual target.

A Khepera equipped with a vision module was placed in a rectangular arena of 160 x 140 cm surrounded by white walls with two black cylinders located in the east and west side of the arena (Fig. 7). The robot has to approach the bigger (B) but not the smaller (S) cylinder. In other words, it has to visually discriminate the two objects. As in the case of the experiment described in the previous section, the robot has to accomplish this task in changing environmental conditions. In particular, the radius of the small and large cylinders were randomly selected each trial between 5 and 125 cm and between 15 and 185 cm respectively.

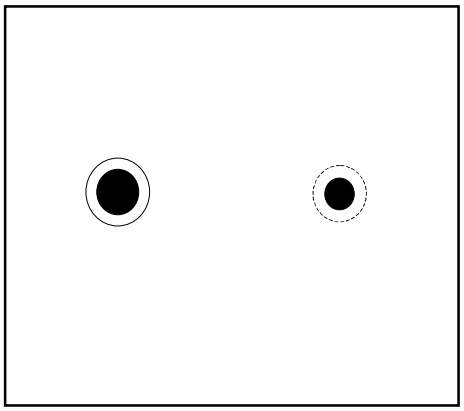


Fig. 7. The environment. The lines represent the white walls. The full circles represent the two cylinders. The plain and dashed circles represent the areas in which the robot is rewarded and punished respectively.

One interesting aspect of this task is that often the robot receives the same sensory patterns while it is facing B or S. In particular B a certain distance looks like S at a closer distance. Therefore, to solve its task, the robot has to overcome the problem that often sensory states are ambiguous. To stress this point, in each trial, the robot is initially place with a random selected orientation along the vertical line from which the two visual objects have the exactly the same perceived size.

The visual system of the robot consists of a linear camera with one array of 64 photoreceptors providing a linear image composed of 64 pixels and subtending a total view-angle of 36°. However, only 8 evenly spaced photoreceptors out of the 64 were used. Evolving individuals have 16 sensory neurons (encoding the state of 8 infrared sensors and of 8 photoreceptors of the camera) and two motor neurons encoding the speed of the two wheels.

If we evolve individuals provided with a simple reactive controller in which the sensory neurons are directly connected to the motor neurons[5] we obtain individuals that are able to

[5] The genotype encoded the connection weights and biases of the neural controller. Each parameter was represented in the genotype by a 8-bit string and normalized between –10 and +10. Population size was 100. The best 20 individuals of each generation were allowed to reproduce by generating 5 copies of their genotype with 2% of their bits replaced with a new randomly selected value. Lifetime consisted of 40 epochs each consisting of a maximum of 200 sensory-motor cycles (each cycle lasted 100 ms). An epoch ended either when the robot reached the rewarding or the punishing area or after 200 sensory-motor cycles (corresponding to 20 seconds at maximum on the physical robot). Individuals' fitness was increased or decreased of 1 point each time the robot reached the rewarding or punishing area respectively. Sensory neurons were updated on the basis of the current state of the sensors and motor neurons were updated according to the standard logistic

correctly navigate toward B in about 60% of the cases on the average and 75% in the case of the best replications of the experiment.

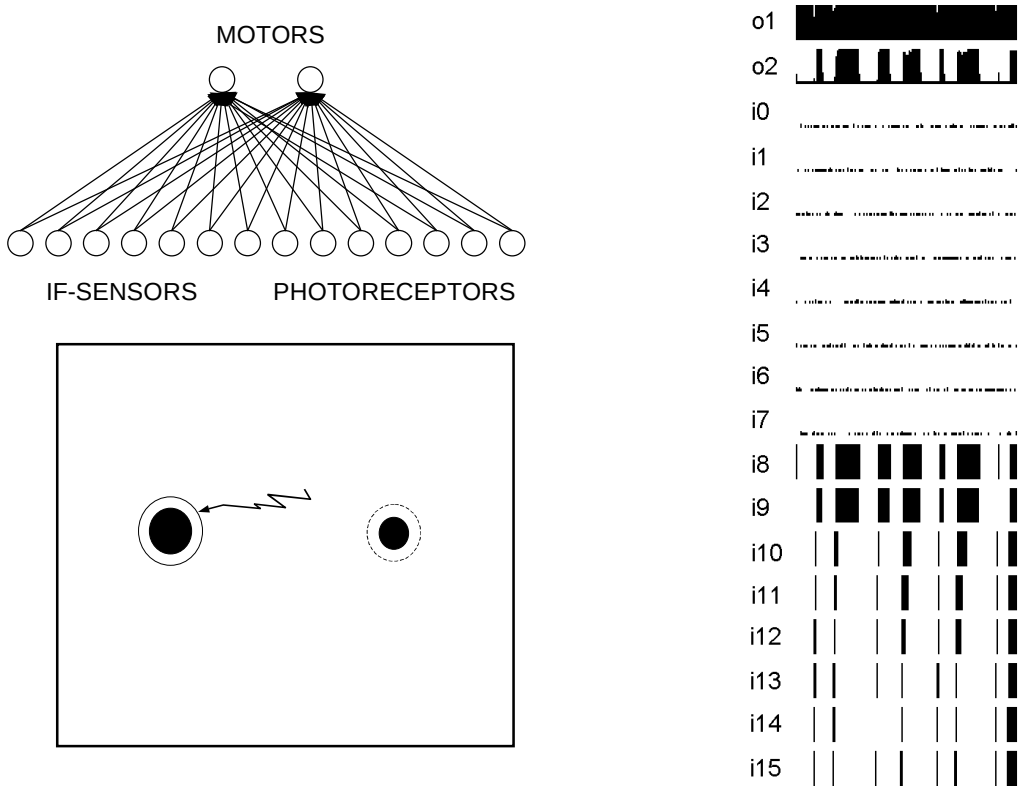


Fig. 8. **Top, left**: the architecture of the neural controller. **Bottom, left**: typical behavior displayed by an evolved robot. **Right**: state of the two motor neurons (o0 and o1) and of 16 sensory neurons (from i0 to i15) throughout time of the individual displayed on the left side of the figure.

As shown in Fig. 8, also in this case evolved individuals strongly rely on sensory-motor coordination to accomplish their task. It should be noted that to navigate toward B, robots should not only take into account the perceived size of the visual target but also its relative distance. Indeed evolving robots act so to loose visual contact after a period of time that is proportional with the distance and the perceived size of the object. This in turn allows evolved robots to move more toward B than toward S independently of the relative size of the two objects. In other words, robots compute the relative sizes of the two objects by exploiting sensory-motor coordination.

This can be illustrated by analyzing the behavior of the individual shown in Fig. 8. The robot rotates clockwise on the spot until it starts to perceive one of the two objects with the camera (in the case, the robot happens to face S first). At this point it starts to move forward by slightly turning clockwise. This allows the robot to loose the visual contact with the object after some time. At this point the robot starts to rotate clockwise again and then to move toward the other object for a while by slightly turning clockwise. The robot thus moves back and forth between the two objects. However, given that it looses visual contact more slowly with the larger and more distant than with smaller and closer objects, it tends to move more toward B than toward S. This in turn allows the robot to finally reach B after a certain

function. The evolutionary process was conducted in simulation by using the Evorobot software (Nolfi, 2000). For an analysis of the performance of the evolved individuals in the physical environment see Nolfi and Marocco (1999).

number of back and forth segments (where each segment consists in a certain number of forward movements toward one of the two objects).

To explain why the robot looses the visual contact more slowly with larger than smaller objects (from the robot position point of view) one should consider that the former persists longer than the latter in the view angle of the camera while the robot move forward by slightly turning clockwise. To explain why the robot loses the visual contact more slowly with more distant than closer object (from the robot point of view) one should consider that, while approaching an object, its angular projection varies more slowly for more distant than for closer objects. In other words, the robot uses motion not only to approach the desired target but also to discriminate between identical sensory patterns corresponding to objects with different size.

The problem with this strategy that explains while it does not allow robots to reach optimal performance is that it must be tuned to the average size of the two landmarks. Indeed, if the robot looses visual contact with the objects too quickly (i.e., if it turns clockwise significantly while approaching an object), it will move more toward B than S overall but it might be unable to reach B within the 200 lifecycles given that it will cover a small distance in every segment. On the contrary, if the robot looses visual contact too slowly, it might risk to reach the wrong object during the first approaching segment without the possibility to measure the relative size of the two objects by moving toward both objects at least one time. Given that how quickly the robot loses visual contact also depends on the absolute size of the objects, the propensity to loose visual contact should be tuned to the average size of the objects themselves. Individuals display a good generalization ability. However, when the range of variation is too high (as in the case of this experiment) evolved individuals might be unable to navigate toward B when it is very large and/or very small.

As in the case of the experiment described in the previous section, individuals controlled by a neural network with two additional internal neurons (receiving connection from the sensory neurons and the internal neurons themselves and projecting connections to the two motor neurons) display much better performance (70% of navigations toward B on the average and 97% in the case of the best replications of the experiment)[6].

Fig. 9 displays the typical behavior exhibited by evolved individuals (see the thick line in the bottom-right of the figure). As in the case of the reactive robot described in Fig 8., this robot rotates clockwise until one of the two objects (S in this case) enter in the view angle of the camera. At this point it starts to move toward the current object by slightly turning clockwise so to loose the visual contact with the object after some time. In this case however, individuals consistently move toward the right target after a while. For instance, the individual shown in Fig. 9 moves toward S only once before moving and reaching B during the second segment. The fact that the same strategy that we described above still play an important role in these individuals can be demonstrated by lesion the internal neurons. In the case of the individual described in Fig. 9, for instance, by freezing the activation state of the two internal neurons to 0, the individual displays exactly the same strategy displayed by the reactive robots described above (i.e., the individual approach B by moving back and forth between B and S and by loosing visual contact more quickly with S than with B). See the dashed line in the bottom-left part of the figure.

[6] The genotype encoded the connection weights and biases of the neural controller and the time constant parameters of sensory and motor neurons. Each parameter was represented in the genotype by a 8-bit string. Weights and biases were normalized between –10 and +10. Time constants were normalized between 0.0 and 1.0. Sensory and internal neurons were updated on the basis of equations (1). Motor neurons were updated according to the standard logistic function (i.e. the time constant parameter was set to 0). All other parameters were the same of the experiment described above.

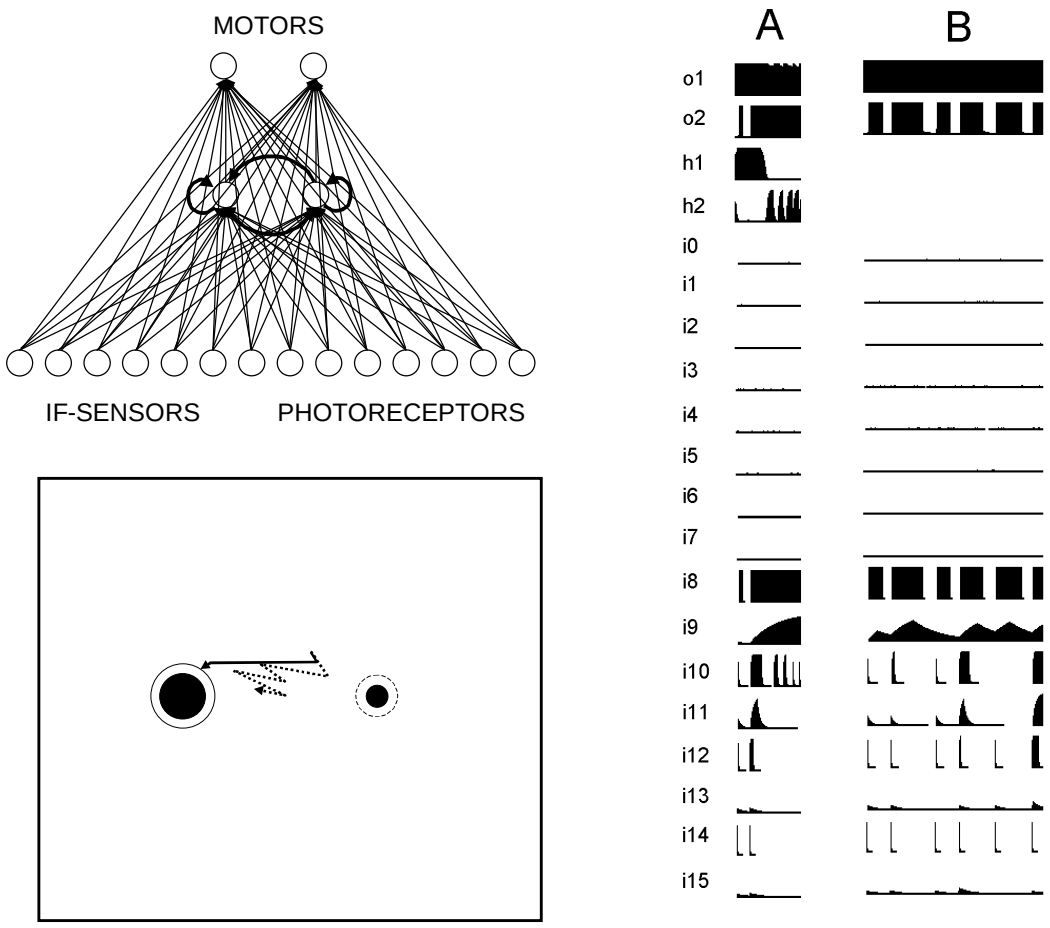


Fig. 9. **Top, left**: the architecture of the neural controller. **Bottom, left**: typical behavior displayed by an evolved robot (the full line represents the trajectory of the robot in normal conditions, the dashed line represents the trajectory produced by the robot when the activation state of the two internal units is set to 0). **Right**: (A) state of the two motor neurons (o0 and o1), two internal neurons (h0 and h1), and of the 16 sensory neurons (from i0 to i15) throughout time of the individual displayed on the left side of the figure. (B) state of the neurons throughout time when the activation state of the two internal neurons is set to 0.

This basic strategy is modulated by the internal neurons so to avoid the problems described above. In particular, in the case of the individual described in the Fig. 9, the second internal neuron modulates the robot behavior so to not loose the visual contact with the current object and consistently move toward it. The first internal neuron instead, by inhibiting the second internal neuron, force the robot to quickly loose visual contact with the current object. During the very first cycles the first internal neurons get fully activated and therefore the robot starts to move toward the current object by quickly loosing the visual contact with the object. This allows the robot to move more toward B than toward S, on the average according to the same strategy that we described above. Later on, however, the first internal neuron tends to be inhibited and, as a consequence, the robot might consistently move toward one of the two targets without going back and forth anymore thus avoiding the risk to not reach B in time. There are several factors that contribute to inhibit the first and activate the second internal neurons. A first factor is simply time (i.e. more is the time that the robot spent negotiating with the two objects higher is the probability that it is now closer to B and higher is the risk to not reach B in time). A second factor is how much time the robot has spent moving toward

the object that is currently in the view angle of the camera (higher is the distance traveled during a single segment higher is the probability that the current object is B). A third factor is the average size of the two landmarks. When both B and S are very small the robot, by loosing the visual contact with the objects too quickly, might risk to not reach B in time. In these situations the second internal neuron becomes slightly activated so to allow the robot to loose the visual contact with the two objects more slowly.

To summarize, as in the experiment described in the previous section, more complex individuals that are able to integrate information over time still rely on the same sensory-motor strategy adopted by reactive individuals and do not discover a completely different strategy. The additional internal states are used to modulate the basic sensory-motor strategy by taking into account the information gathered during the previous interaction with the environment (e.g., the time spent travelling toward the object which is currently in the view angle of the robot) and the current environmental characteristics (e.g., the average size of the two current objects).

8 Conclusions

In this paper we discussed in which conditions we can expect the emergence of agents able to integrate sensory-motor information over time and later use this information to modulate their behavior accordingly.

We showed how the border between what can be accomplished by simple agents that only rely on their current sensory states or on their internal dynamics and by more complex agents that are also able to integrate information over time is rather fuzzy and cannot be formally identified. However, we showed how problems that should be accomplished in varying environmental conditions tend to require agents able to integrate information over time.

By evolving robots for their ability to solve a navigation and a visual discrimination task in changing environmental conditions we showed how evolved individuals provided with a suitable neural architecture that allows them to integrate sensory-motor information over time display optimal or close to optimal performance and significantly outperform individuals that can only rely on simpler reactive strategies. On the other hand we showed that the former individuals that are able to extract information overt time also rely on the same reactive strategies discovered by the latter individuals. Such simple reactive strategies, that exploit sensory-motor coordination, are complemented and enhanced with an ability to integrate information over time and later use such information to modulate the robot's behavior. More complex and effective individuals indeed looks organized in a hierarchical fashion where higher levels, that also rely on internal states, modulate the activity of lower layers, that only rely on simple reactive mechanisms.

The fact that mixed strategies of this sort exist is important from an evolutionary point of view because evolving individuals might progressively increase their ability to develop and to rely on internal mechanisms, in addition that on sensory-motor coordination, without necessarily undergoing a profound re-organization. In other words, these mixed strategies might represent a sort of stepping stones toward significantly more complex strategies that heavily rely on the ability to integrate sensory-motor information over time and to modulate the later behavior accordingly.

The analysis of evolved individuals also revealed that internal states can be more properly described as partial, action-oriented, and action-mediated rather than as complete, action-independent, and objective representation of the external environment.

That internal states do not necessarily had to include a complete and explicit representation of the external world have been claimed by Brooks (1991) among many others. According to Brooks, representations can be partial models of the world and include only the aspects that are relevant for achieving a given task. In addition representations can be action-oriented, i.e. the form that these representation take depends on the particular task those representations are to be used for. Indeed the strategy adopted by evolving individuals to solve the problems that we described in this paper resort on representations that are local and action-oriented. For example, in the case of the experiment described in section 7, individuals have internal states that apparently encode how they should act (i.e., if the have to keep traveling toward the current object or not) and not objective information about the external environment (e.g. which is the size of the current object).

Moreover, the analysis of the evolved individuals suggests that representations tend to be action-mediated, i.e., they implicitly and indirectly encode information about the external environment and/or the agent/environment relation by encoding information on what the agent did in its previous interaction with the environment. For example, in the case of the experiments described in section 7, evolved individuals indirectly and implicitly encode information about the relative size of the current object by taking into account how much time they spent travelling toward it.

Whether the possibility to generate an open ended evolutionary process able to sintesize more and more complex individuals able to solve a large variety of complex tasks might lead to individuals that will mostly rely on action-independent and objective representations or whether action-dependent and action-mediated representations will still play an important role is an open question. It should be noted, however, that action-mediated representations might be preferred not only because they are simpler and easier to synthesize from an evolutionary point of view but also because they are less affected by the problems described in section 5. In particular they might be less affected by problems of reliability (this can be explained by considering that, in general, proprioceptive information encoding what the robot is doing is less affected by unpredictable causes such as ambient light or noise that heavily affect information encoding how the environment looks like). Moreover, action-mediated representation might be less affected by problems deriving from an imperfect coupling between the internal states and the sensory-motor flow (this can be explained by considering that proprioceptive information is immediately coupled with the robot behavior while objective information, such as a map of the external environment, should be continuously coupled with the changes in position and orientation of the robot).

References

Ackley D.H. and Littman M.L. (1991). Interaction between learning and evolution. In C.G. Langton et. al (Eds.), *Proceedings of the Second Conference on Artificial Life*. Reading, MA: Addison-Wesley.

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

Beer R.D. (1995b). Computational and dynamical languages for autonomous agents. In R. Port and T. van Gelder (Eds,), Mind as Motion. Cambridge, MA: MIT Press.

Beer R.D. (1996). Toward the evolution of dynamical neural networks for minimally cognitive behavior. In P. Maes, M. Mataric, J. Meyer, J. Pollack and S. Wilson (Eds.), *From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*. Cambridge, MA: MIT Press.

Beer R.D. and Gallagher J. C. (1992) Evolving dynamical neural networks for adaptive behavior. *Adaptive Behavior*, 1:91-122

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.

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

Elman J.L. (1990). Finding structure in time. *Cognitive Science*, 14:179-211.

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

Floreano D. and Mondada F. (1996b) Evolution of plastic neurocontrollers for situated agents. In P. Maes, M. Mataric, J-A. Meyer, J. Pollack & S. Wilson. (Eds.), *From Animals to Animats 4, Proceedings of the International Conference on Simulation of Adaptive Behavior.* Cambridge, MA: MIT Press.

Gallagher J.C. and Beer, R.D, Espenschied, K.S. and Quinn, R.D. (1996). Application of evolved locomotion controllers to a hexapod robot. *Robotics and Autonomous Systems* 19:95-103

Gallistel C.R. (1990) *The Organization of Learning*. Cambridge, MA: MIT Press.

Gomi T. & Ide K. (1998) Emergence of gaits of a legged robot by collaboration through evolution. In P. K. Simpson (Ed.), *IEEE World Congress on Computational Intelligence (WCCI'98)*. New York: IEEE Press.

Harvey I., Husband I. and Cliff D. (1994) Seeing the light: artificial evolution, real vision. In D. Cliff, P. Husband, J-A Meyer & S. Wilson (Eds.), *From Animals to Animats 3*, *Proceedings of the Third International Conference on Simulation of Adaptive Behavior*. Cambridge, MA: MIT Press/Bradford Books.

Hendriks-Jansen H. (1996). *Catching Ourselves in the Act.* Cambridge, MA:MIT Press.

Husband P., Smith T., Jakobi N. & O'Schea M. (1999). Better living through chemistry: Evolving GasNets for robot control. *Connection Science*, (10)3-4:185-210.

Jakobi N. (1998) Running across the reality gap: Octopod locomotion evolved in a minimal simulation. In P. Husbands and J-A. Meyer (Eds.), *Evolutionary Robotics. First European Workshop*, pp.39-58. Berlin: Springer-Verlag.

Jordan M.I. (1989). Serial Order: A parallel, distributed processing approach. In J. L. Elman & D. E. Rumelhart (Eds.), *Advances in Connectionist Theory*. Hillsdale, NJ: Erlbaum.

Lewis M.A., Fagg A.H. & Solidum A. (1992) Genetic programming approach to the construction of a neural network for control of a walking robot. In: *Proceedings of the IEEE International Conference on Robotics and Automation*, pp.2618-2623. New York: IEEE Press.

Mayley G. (1997). Landscapes, learning costs, and genetic assimilation. *Evolutionary Computation*, 4(3):213-234.

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

Nolfi S. (1997a) Using emergent modularity to develop control system for mobile robots. *Adaptive Behavior*, 3-4:343-364.

Nolfi S. (1997b) Evolving non-trivial behaviors on real robots: A garbage collecting robot. *Robotics and Autonomous System*, 22:187-198.

Nolfi S. (2000). Evorobot 1.1 User Manual. *Technical Report*, Institute of Psychology, Rome (available at http://gral.ip.rm.cnr.it/evorobot/simulator.html).

Nolfi S. (in press). Power and limits of reactive agents. *Neurocomputing*.

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

Nolfi S., Marocco D. (2000). Evolving visually-guided robots able to discriminate between different landmarks. In: J-A Meyer, A. Berthoz, D. Floreano, H.L. Roitblat, and S.W. Wilson (eds.) *From Animals to Animats 6. Proceedings of the VI International Conference on Simulation of Adaptive Behavior*. Cambridge, MA: MIT Press.

Nolfi S. & Marocco D. (in press). Evolving robots able to integrate sensory-motor information over time. In R.J. Duro, J. Santos, M. Grana (eds.) Biologically Inspired Robot Behavior Engineering. Heidelberg, Germany: Springer Verlag.

Nolfi S. and Parisi D. (1993). Auto-teaching: networks that develop their own teaching input. In J.L. Deneubourg, H. Bersini, S. Goss, G. Nicolis and R. Dagonnier (Eds), *Proceedings of the Second European Conference on Artificial Life*. Brussels: University Libre of Brussels.

Nolfi S. and Parisi D. (1999). Exploiting the power of sensory-motor coordination. In: D. Floreano, J-D. Nicoud, and F. Mondada (Eds.). *Advances in Artificial Life, Proceedings of Firth European Conference on Artificial Life*, Berlin: Springer Verlag.

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

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

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

Whitehead S.D. and Ballard D.H. (1991) Learning to perceive and act by trial and error. *Machine Learning* 7:45-83.