Simulating exploratory behavior in evolving Artificial Neural Networks

Richard Walker
Università di Palermo
Viale delle Scienze
Palermo
Italy
rwalker3@csi.com

Orazio Miglino
Consiglio Nazionale delle Ricerche
Viale Marx, 15
00145 Roma
Italy
orazio@caio.irmkant.rm.cnr.it

Abstract

Animals released into unfamiliar environments will often engage in "roaming" behavior, apparently for exploratory purposes. It is likely that this behavior constitutes a "behavioral primitive" which can be used in the construction of more complex behaviors. This paper reports a series of experiments in which a Genetic Algorithm is successfully used to "evolve" efficient exploration strategies in a population of software simulated Khepera robots, controlled by Artificial Neural Networks. Robots based on simple perceptrons with no hidden neurons outperformed those with more complex control networks. These robots tended however to adapt to the specific environments where they had evolved. More robust behavior was obtained from robots where input from the external environment was enriched with data from cyclical "time sensors". It is suggested that control-networks based on this architecture could become a useful component in more complex systems.

1 INTRODUCTION

When a mouse is released into an unfamiliar environment it will "wander" or "roam" in an apparently random fashion. If we track the animal's path during roaming we discover that the exploration is "efficient" - the mouse rapidly covers a large proportion of the available territory (Carr & Watson, 1908). Similar behavior may in fact be found in a broad variety of species (Gallistel, 1990). It seems very likely that the function of this behavior is exploratory: "roaming" or "wandering" enables the animal to learn about the new environment, to locate sources of food or danger, to create an internal map of the terrain or to identify landmarks which can be used in subsequent navigation (Thinus-Blanc, 1996). Wandering may also be a low-level building block used in the construction of more complex behaviors. (e.g. reaching a target while performing a "detour" around an obstacle (Corbacho & Arbib, 1995)). An ability to explore unknown environments could be useful for man-made autonomous or semi-autonomous agents as well as for animals. A robot engaged in planetary exploration could search a terrain for interesting geological material without waiting for commands from earth; a software agent exploring the World Wide Web could automatically decide which site to visit next. Investigation and simulation of animal exploration strategies is thus of engineering as well as scientific interest.

The problem of how to efficiently explore an environment of unknown shape and configuration is computationally equivalent to the task of "filling" a polygon in computer graphics. The problem can be easily solved by the use of recursive techniques (e.g. "flood-fill" or "boundary fill" algorithms). It should however be noted that these algorithms require the computational system to keep track of the areas it has already visited. This approach is memory intensive. Cognitive science models positing the existence of detailed topological maps suffer from the same difficulty. Perhaps more seriously the sensory input available to robots or biological organisms, operating in real-world environments, may not be rich enough to generate high resolution location information. It is interesting therefore to investigate the feasibility of efficient exploration without the use of topological maps or similar representations.

2 RELATED WORK

"Wandering"and "roaming" by animals has been studied extensively by field biologists (Gallistel, 1990). Exploratory behavior has often been reported in laboratory experiments designed to explore more complex forms of behavior (Carr & Watson, 1908). It seems likely that "wandering" behavior is closely related to foraging. According to "Optimal foraging theory" animals forage "efficiently", maximizing benefits and minimizing costs (Stephens & Crebs, 1986). Other workers in the field argue, on the other hand, that actual animal behavior may often be severely sub-optimal, owing to cognitive limitations and poor perceptual input (Gould & Lewontin 1979). The debate is an interesting one with implications for general issues in theoretical biology. To the knowledge of the authors there have, however, been relatively few attempts, prior to the work reported in this paper, to model the basic computational mechanisms underlying this kind of behavior - perhaps considered too simple to deserve serious attention.

The use of simulated robots to explore cognitive mechanisms in animals was first introduced (at a conceptual level) by V. Braitenberg (Braitenberg 1984). Arbib (Arbib, 1987) has used simulated (hand-designed) robots to investigate animal exploration based on visual clues. Dorigo and Colombetti (Dorigo & Colombetti, 1994) applied evolutionary techniques in the design of physical robots with the ability to perform simple behavioral tasks (light approaching etc.). Wee Kheng Leow (Kheng Leow, 1998) has used evolutionary techniques to investigate exploratory behavior on simulated robots guided by smell

The general methodology used in our experiments is inspired by the basic techniques developed by Nolfi, Floreano, Miglino and Mondada. (Nolfi, Floreano et al., 1994). The simulator software used in the experiments was developed in previously published work by Miglino, Lund and Nolfi (Miglino, Lund et al., 1996). In this work the robot was trained to perform an obstacle-avoidance task. Other work based on the same basic model has included the modelling of other forms of obstacle avoidance (Nolfi, Miglino et al., 1994) as well as a "garbage collection" task (Nolfi & Parisi 1997) and detour behavior (Miglino, Denaro et al., 1998).

"Wandering" behavior has been investigated by Miglino, Nafasi and Taylor (Miglino, Nafasi et al.,1995) who used techniques derived from evolutionary robotics to develop an ANN-controlled mobile Lego robot trained to explore an open arena.

3 OBJECTIVES AND METHODS

The aim of our investigation was to find the simplest possible computational mechanism capable of generating efficient exploratory behavior, if possible, without resort to detailed tracking information or topological maps. In order to achieve this goal we used a Genetic Algorithm to evolve robots exhibiting the desired behavior.

The genetic algorithm operated on a population of 100 robots, simulated in software. The basic structure and behavior of the robots was based on the Khepera robot, developed at E.P.F.L. Lausanne (Mondada, Franzi et al., 1993).The simulator software was based on careful measurement of the behavior of the physical robot in real life environments (Miglino, Lund et al., 1996).

Khepera has a circular shape with a diameter of 5.5 cm. The robot's basic architecture is based on three main components: a set of sensors, a motor apparatus and a control network. In the experiments reported here the input to the control network comes from 8 infrared proximity sensors, two of which are positioned on the front of the robot, two on the back and two on each side. Each proximity sensor is sensitive to obstacles with a range of 3 cm and has an angle of vision of 20 degrees. Output values are computed as a function of the position of the robot, using data matrices collected from Khepera. In one experiment input from the proximity sensors was enriched with input from two "internal clocks" The activation value of a time sensor at time t is computed using the following algorithm:

$IF\,(activation_t < 1)\,THEN$

$$activation_t = activation_{t-1} + k$$

$ELSE$

$$activation_t = 0$$

For the Khepera control network we used 2 different time sensors with $k = 0.002$, $k = 0.004$.

The motor apparatus consists of a left and a right wheel driven by stepping motors which can move both forwards and backwards. With both stepping motors working at full speed the robot is capable of moving 10 cm/cycle of computation. The Control Network is an Artificial Neural Network (ANN) whose input neurons represent the state of the sensors and whose output neurons control the stepping motors. Neuron activation levels are a logistic function of total input to the neuron including input from a permanently active bias neuron.

In our experiments we tested three different network topologies: a fully connected perceptron (Minsky & Pappert 1988) with eight infrared sensors

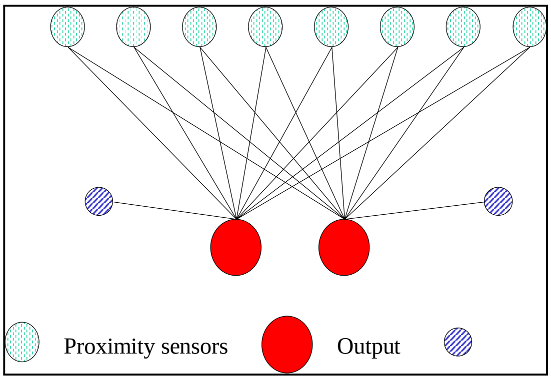

Figure 1: Fully connected perceptron

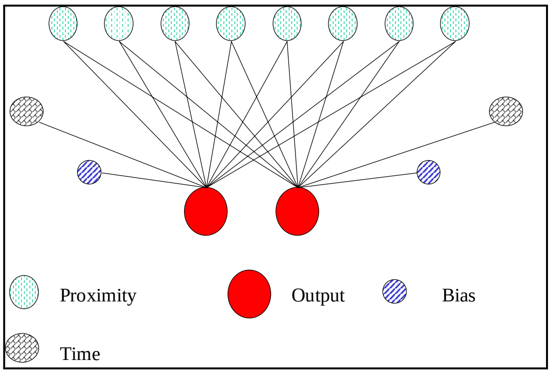

Figure 2: Perceptron with time sensors

(see Figure 1), a second fully connected perceptron with 8 infrared sensors and two internal clocks (Figure 2) and a fully-connected 3 layer feed-forward network with 5 hidden neurons (Figure 3).

Evolution involved "mutations" in the strengths of the connections linking the output neurons to the input and bias neurons. The genome of the organism consisted of a sequence of binary coded numbers (8 bits per number) representing the strengths of individual connections.

In our experiments we evolved efficient exploration strategies by applying artificial selection to the population of control networks. In the initial population connection strengths for individual networks were set to random real values uniformly distributed between -1 and 1. Each robot was placed in an "open field box" (a square or rectangular terrain surrounded by a "fence"). Robots were tested four times in a rectangular 90cm by 40 cm box and four times in a square 60 cm by 60 cm box. On each test the robot started with a randomly chosen position and orientation. For measurement purposes the terrain was divided into square

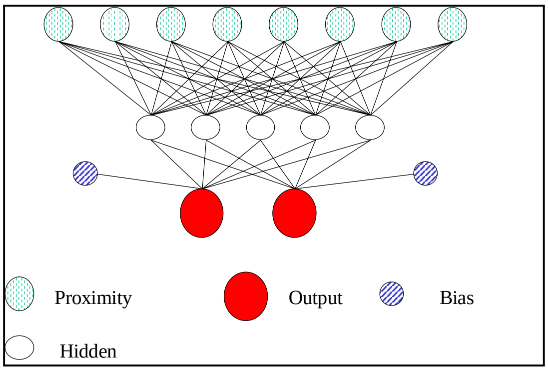

Figure 3: Perceptron with hidden units

10 cm by 10 cm cells. Each test consisted of 1,000 cycles of computation. Efficiency of exploration was measured by a fitness function which rewarded completeness and speed of exploration.

$fit = ncells + (nCycles - cyclesToCompletion)$

Where:

fit is the fitness attributed to the robot

$nCells$ is the number of different cells touched by the robot during exploraiton

$nCycles$ is the duration of the exploration (in cycles)

and

$cyclesToCompletion$ is the number of cycles traversed before the robot has touched every cell in the box. (If the robot never touches all the cells $cyclesToCompletion$ is assigned the same value as $nCycles$.)

It will be noted that the fitness function used is additive rather than multiplicative - this ensures that any increase in the number of cells traversed during exploration will translate into improved fitness even if the robot does not succeed in visiting all the cells in the box.

When all robots had been tested the 20 robots with the highest fitness scores (summed over the eight tests) were selected for "reproduction". Each of the selected robots produced 5 offspring. Reproduction was asexual. During the cloning process "mutations" were introduced by flipping bits in the genome with a probability of 0.04 per bit per generation. This process was iterated for 100 generations by which time no further improvements in fitness were observed.

In order to guarantee the statistical robustness of our results each simulation was repeated five times using a different random number seed on each occasion.

Finally, the robots produced in each of the simulations were tested in a new environment (a 120 cm by 90 cm box) which they had not encountered during the evolutionary process. This test made it possible to measure the robustness of evolved network configurations with respect to changes in the shape of the environment to be explored. As in previous tests each robot was tested from 4 different starting positions. Given the larger size of the new environment the number of computation cycles was increased to 3,000.

4 RESULTS

4.1 OVERALL FITNESS RESULTS

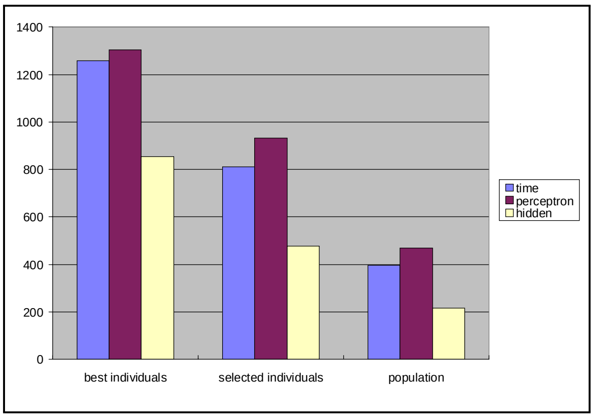

Figure 4: Mean fitness on last generation

Figure 4 shows the mean levels of fitness (over the five simulations) achieved by individuals in the last generation tested. Data is presented for the whole population, the population of individuals selected for reproduction and the best individual in the population. Analysis of variance shows that for each of these groupings the architecture with hidden neurons - the most complex of those tested - achieved significantly lower levels of fitness than the combined results for perceptron and time sensor architectures (PopulationF(1498,1)=4.89, $p<0.05$; Selected:F(298,1)=4.35, $p<0.05$; Best:F(73,1)=4.91, $p<0.05$).

Limiting the comparison to the Time Sensor and the Perceptron-based architectures the Perceptron achieved significantly higher levels of fitness in the population as a whole (F(1498,1)=3.96, $p<0.05$) and in the group selected for reproduction (F298,1)=4.09, $p<0.05$). There was however no statistically significant difference between the results achieved by the best organisms with these two architectures (F(73,1)=3,45, $p>0.05$).

In terms of robustness (the ability to maintain their efficiency in unfamiliar environments) the fittest robots were again those based on perceptron and time sensor networks. For clarity of presentation we will limit our statistical analysis to a comparison between the average performance for robots with these architectures.

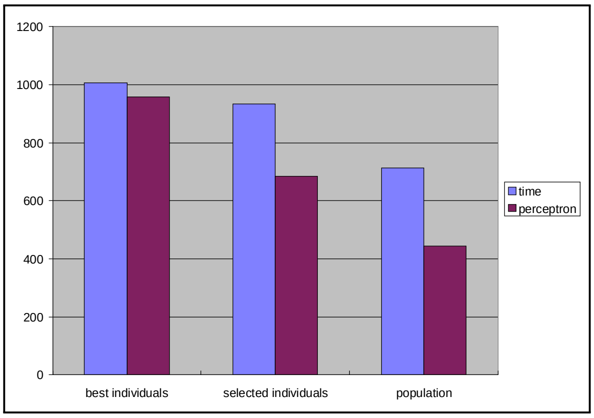

Figure 5: Mean performance in unfamiliar environments

Figure 5 compares the two architectures in terms of average performance for the whole population, robots selected for reproduction and the best performing individual. In each of these comparisons the time-sensor based robots achieve significantly higher performance than the perceptron models. (population: t(998)=2560, $p<0.01$; selected: t(198)=2240, $p<0.01$; best: t(48)=2100, $p<0.01$).

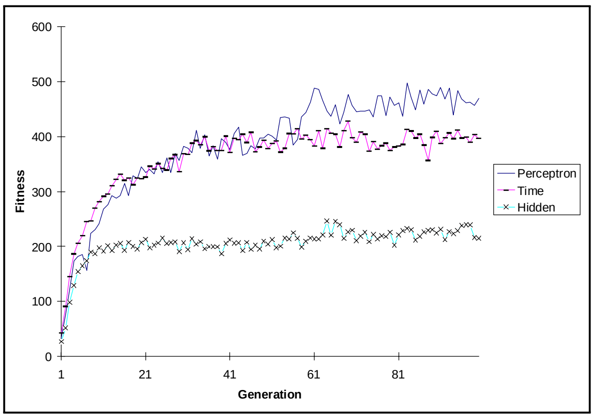

Figure 6: Evolution of fitness over time

Figure 6 plots the evolution over time of mean population fitness for the three network architectures tested. From this data it appears that high levels of fitness are associated with a longer evolutionary process. The hidden neuron architecture - the least efficient of those tested - reached its fitness plateau after approximately 20 generations; the neurons with time sensors failed to improve their fitness after the fortieth generation; robots based on the Perceptron architecture continued to improve their fitness up to the eightieth generation.

4.2 CHOICE OF EXPLORATION STRATEGIES

Figures 7 and 8 show examples of the trajectories followed by fully evolved Perceptrons while exploring a rectangular open field box. These can be compared with the trajectories for time sensor networks presented in Figure 9. A qualitative examination of these trajectories facilitates the explanation of the fitness data presented earlier.

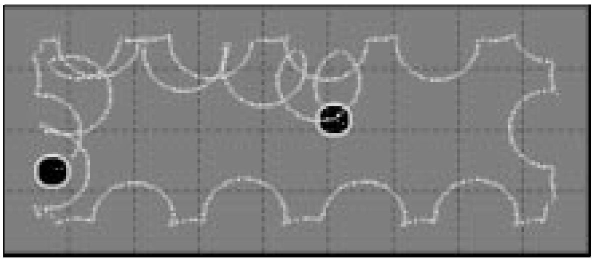

Figure 7: Trajectory of a Perceptron-controlled robot

The perceptron shown in Fig 7 followed a trajectory consisting of a sequence of alternating semi-circle and straight line segments. In the rectangular field box this strategy enabled the robot to rapidly achieve complete cell coverage. The high fitness score achieved in this environment is more than enough to compensate the relatively lower score achieved in the square box where the strategy adopted failed to cover cells in the middle of the field box (trajectory not shown).

The second perceptron whose trajectory is shown in Fig 8 has evolved a simple strategy which consists of proceeding in a straight line and turning by roughly 90 whenever the robot meets an obstacle.

For both robots the strategy which emerges during the evolutionary process is highly adapted to the specific environment where the robot has evolved. In the case of the first robot the diameter of the semi-circles and the length of the line segments in the trajectory are perfectly adapted to to the cell-size used for evaluation purposes. In the second case the ability of the evolved strategy to cover all the cells in the box depends on the

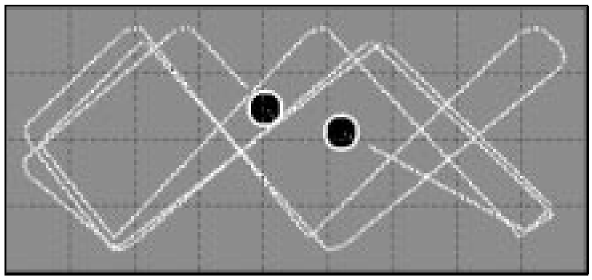

Figure 8: Trajectory of a second Perceptron-controlled robot

precise relationship between the length and the width of the field box.

The easy to describe, environment-dependent strategies evolved by perceptron-based robots may be compared with the more complex behavior of robots with time-sensors (see Figure 9).

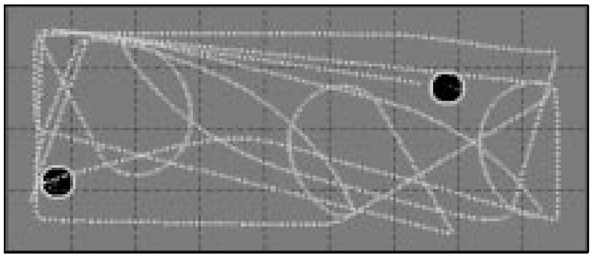

Figure 9: Trajectory of a robot with time sensors

In the experiments with these robots we observe, as in the previous case, that the robot turns on meeting an obstacle. It appears however that the robot also turns (at an oblique angle) on receiving input from the time sensors. In many cases the robot will then traverse the field box following a curved trajectory which maximizes the number of cells visited en route. The ability to periodically change direction is independent of the external environment where the robot has evolved.

Comparing the perceptron and the time-sensor based robots it may be observed that in the absence of any other source of input the former have necessarily adapted to the specific environments where they evolved. The latter on the other hand have access to a source of input (the time sensors) whose characteristics are independent of the external environment. This provides a possible explanation for the superior performance of time-sensor based robots in unfamiliar environments (see Figure 10)

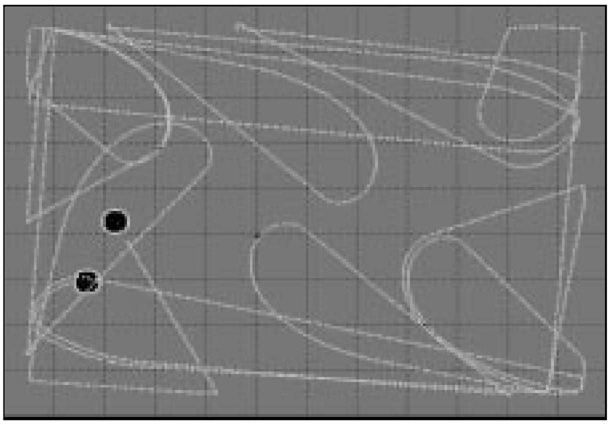

Figure 10: Trajectory of a time sensor robot in an unfamiliar environment

5 DISCUSSION

The results of the experiments presented in this paper show that Genetic Algorithms can be used to generate effective exploratory strategies for autonomous agents. It shows, furthermore, that the minimal computational mechanism capable of generating such behavior is extremely simple. In the perceptron and time-sensor experiments the control network for an individual was completely specified in 18 and 22 bytes respectively. Networks with hidden neurons (specified in 52 bytes) were significantly less efficient than the simpler models. The population used - 100 individuals - was extremely small. Even the most efficient architectures achieved a fitness plateau in around 80 generations. The combination of simple architecture, small population size and a limited number of generations meant that compute time was limited. A complete set of simulations for a particular architecture can be computed on a Pentium 200 Personal Computer in less than two hours.

The results achieved have both cognitive science and engineering implications. From the former viewpoint it is interesting to note how extremely simple neural networks can generate highly efficient strategies of exploration. It is clear that the quantity of memory used to specify each control network is insufficient to contain a “map” of the open field box. It thus appears that at least for purposes of exploration (if not for more complex tasks) high resolution location data (stored as a topological map or in other forms) is superfluous. In more general terms this suggests that cognitive scientists and network designers should be cautious in their assumptions with respect to the complexity of the networks needed to achieve specific goals. The results presented here confirm that in some cases complex networks with hidden neurons may be not only unnecessary but actually counter-productive.

A second interesting result of the experiments is the superior generalization ability of networks with time-sensors. Animal psychologists have long recognized the importance of internal “clocks”, in animal behavior. To the knowledge of the authors, however, this is the first time that this kind of sensor has been used in evolutionary robotics. At this stage of our research the exact role the time-sensors are playing is not entirely clear. What is evident however is that if behavior is made to depend exclusively on stimuli from the outside world (as in our perceptron-based robots) natural selection will inevitably lead to the evolution of highly specialized solutions with only limited ability to adapt to environmental change. The results of our experiments suggest that stimuli from internal clocks may play an important role in maintaining the stability and the viability of the organism in the face of changes in the external environment. This result is worthy of note both for the cognitive scientist and the engineer.

6 CONCLUSIONS - DIRECTIONS FOR FUTURE RESEARCH

Much current research in the field of “evolutionary robotics” is based on a “bottom up” approach (Clark 1997). Complex architectures are built up step by step out of simpler modules - complex behaviors are constructed from simple behavioral primitives.

The existence of “wandering” and “roaming” in a broad range of animal species, living in many different environments, is evidence that ”exploration” should be considered a behavioral primitive in biological organisms and suggests that it would be useful to incorporate such behavior in artificial autonomous agents. This would however require the development of control networks capable of performing tasks which were not tested in the current work.

In the research reported in this paper the environments in which robots were tested were convex, highly regular spaces, containing no obstacles. It is evident that to be useful in a real-life context robots should be capable of negotiating non-convex spaces characterized by the presence of irregular boundaries and obstacles. Only in this case could the robot be considered as computationally equivalent to the computer graphics algorithms mentioned at the beginning of this paper.

The authors predict that while “externally driven” devices such as the Perceptron might adapt to the requirements of (some) specific environments it will not be possible for this kind of device to evolve strategies applicable to unfamiliar environments. We believe however that the "internally driven" strategies evolved by "time-sensor" networks will prove robust, even in environments considerably more complex than those where it has been tested up to now. Future experiments will test this hypothesis, hopefully producing robust networks which can be incorporated as sub-systems in the larger, more complex machines which the evolutionary robotics community is currently working to build.

References

Arbib, M.A. (1987), Levels of Modelling of visually guided behavior, *Behavioral and Brain Sciences*, 10, 407-465

Braitenberg, V. (1984), Vehicles, Experiments in Synthetic Psychology, MIT Press

Carr, H., Watson J.B. (1908), Orientation of the White Rat, Journal of Comparative *Neurology and Psychology.* 18, 27-44

Clarke A. (1997), Being There, MIT Press

Corbacho, F.J., Arbib M.A. (1995), Learning to detour. *Adaptive Behavior.* 3(4), 419-468

Dorigo M, Colombetti. M. ,Robot Shaping (1994): Developing Autonomous Agents through Learning. *Artificial Intelligence*, 71(2):321–370,.

Gallistel.C.R. (1990), The Organization of Learning, MIT Press

Gould, S. J & Lewontin, R. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society of London*, Series B 205(1161), 581-598.

Jakobi, N., Husbands P., & Harvey I (1996). Noise and the Reality Gap. The use of simulation in Evolutionary Robotics, *Proceedings of the Third European Conference on Artificial Life (ECAL '95)* - Springer Verlag

Kheng Leow, W. (1998), Computational Studies of Exploration by Smell, *Adaptive Behavior*, 6 (3/4), 411-434

Miglino O., Nafasi K. & Taylor C. (1995), Selection for wandering behavior in a small robot, *Artificial Life*, 1995 2-1

Miglino O., Lund H. H., Nolfi S. (1996), Evolving Mobile Robots in Simulated and Real Environments, *Artificial Life*, 2(4): 417-434

Miglino, O., Denaro, D., Tascini, E., Parisi, D. (1998). Detour behavior in evolving robots. *Proceedings of First European Workshop on Evolutionary Robotics.*

Minsky M. & Papert S., (1988) Perceptrons, MIT. Press

Mondada F., Franzi E. & Ienne P. (1993) Mobile Robot miniaturisation. A tool for investigation in control algorithms. In *Proceedings of the Third International Symposium on Experimental Robotics*, Kyoto, Japan

Nolfi, S., Floreano D., Miglino O., Mondada F., (1994), How to evolve autonomous robots: different approaches in evolutionary robotics. In Brooks R.A. & Maes P. (eds), *Artificial Life IV- Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, pp. 190-197, Cambridge MA, MIT Press

Nolfi, S., Miglino O. & Parisi D. (1994) Phenotypic plasticity in evolving neural networks. In D.P. Gaussier and J-D Nicoud (eds), *Proceedings of the Intl. Conf. From Perception to Action, Los Alamitos*, CA, IEEE Press

Nolfi, S., and Parisi, D., (1997) Evolving non-trivial behaviors on real robots: a garbage collecting robot, *Robotics and Autonomous Systems*

Stephens, D. W. & Krebs, J. R. (1986). Foraging theory . Princeton, NJ: Princeton University Press.

Thinus-Blanc,(1996) C., Animal Spatial Cognition, World Scientific, Singapore