

Social Facilitation on the Development of Foraging Behaviors in a Population of Autonomous Robots

Alberto Acerbi, Davide Marocco, and Stefano Nolfi

Institute of Cognitive Science and Technologies, National Research Council,
Via San Martino della Battaglia 44, Rome, 00185, Italy
{alberto.acerbi, davide.marocco, stefano.nolfi}@istc.cnr.it

Abstract. In this paper we propose an adaptive algorithm based on a combination of selective reproduction, individual learning, and social learning. Social learning consists of a simple facilitation process that regulates the strength of individual learning on the basis of the number of individuals located nearby. By testing this model in an experimental scenario, in which a population of 10 mobile robots has to develop a simple foraging behavior, we demonstrate how the model proposed produces effective results. By comparing the results obtained in different experimental conditions we also show how the method proposed outperforms other alternative algorithms based on genetic evolution or individual learning. Finally, we briefly discuss how the model proposed can help us to understand the role of social learning in biological organisms.

Key words: Social Learning, Social Enhancement, Learning and Evolution.

1 Introduction

It is a widespread opinion that social learning dynamics can account both for the diversity of humans' behavioral repertoire [12] and for the complexity of humans' cognitive skills [13]. This consideration has led most of the researchers that deal with social learning to focus on high-level, human-like, forms of social learning and to attempt to reproduce them through specific algorithms that aim to replicate a behavior of a model by explicitly copying it.

However, when considering learning as a set of adaptive modifications that take place within each single agent in response to environmental stimuli [9], we have to consider that the environment of a single agent also includes its conspecifics. In this perspective, their behavior becomes a valuable source of information that can be exploited in several ways besides from explicit copying. Indeed, especially in the last twenty years, ethologists have shown how other forms of social learning that are much simpler and less cognitive demanding than explicit imitation play an important role in the development of complex behaviors in various vertebrate species. For example, the food preferences developed by Norway rats (*Rattus norvegicus*) are influenced by the food eaten by conspecifics through indirect information obtained by sniffing conspecifics' breath [4]. Similarly, female guppies (*Poecilia reticulata*) prefer to mate with males that they have seen mating before [3]. Again, the development of several primates' behaviors, which have been previously interpreted as a result of explicit imitation, are now considered as an outcome of more simple forms of social

learning like *emulation* or *stimulus enhancement* (for reviews on social learning in animals, but also in simulated agents and robots see [2][5][6][14]).

The development of artificial organisms (robots or simulated agents) through artificial life techniques is an excellent way of exploring the feasibility of simple models of social learning [7][8] with particular reference to the possibility to take into account different aspects of the individuals (their body, their nervous system, and the way they adapt to the environment), of the population, and of their interaction. As formerly pointed out in [8], identifying alternative models of social learning, besides imitative learning, is interesting for both scientific and technological reasons. From a scientific point of view, identifying how a population of artificial agents can obtain adaptive advantages by dealing with simple forms of social learning can shed light on how the interaction between individual learning, social interactions, and evolution might produce effective and robust collective dynamics. Indeed, understanding the interaction between these processes through empirical and laboratory experiments can be very difficult. From a technological point of view, on the other hand, the identification of algorithms that allow embodied agents to adapt by exploiting the interaction between evolution, individual learning, and social learning can lead to the development of innovative methods for developing multi-agent systems, spanning from autonomous robots to ubiquitous computing devices.

In this paper we demonstrate how a simple form of social learning, in which agents learn individually but in which the strength of individual learning is modulated by the number of individuals located nearby, can lead to the development of effective behaviors. This will be demonstrated in an experimental scenario in which a population of simulated robotic agents should develop an ability to forage by discriminating different type of foraging areas. Moreover, we show how, in this type of experimental scenarios, the combination of selection at populational level, individual learning, and social learning outperforms other adaptive processes such as individual learning and genetic evolution.

In section 2 we will describe the experimental setup and the different experimental conditions. In section 3 we show how the results obtained in the social learning conditions outperform those obtained in the individual learning and genetic evolution conditions. Finally, in section 4, we discuss the implications.

2 Experimental Setup

A team of 10 e-puck robots is placed in an environment that consists in a square arena of 200 x 200 cm surrounded by walls. The floor of the arena is grey and contains two circular target areas with a diameter of 60 cm colored in black and white, respectively (Fig. 1). Robots are provided with simple sensory-motor capabilities that allow them to move, gather information from the environment and to produce sound signals with a fixed intensity.

The robots (Fig. 2 Left) have a circular body with a radius of 37 mm, 8 infrared sensors placed around the body, 1 ground sensor placed on the bottom of robot, 1 microphone and 1 speaker. Robots' neural controllers consist of neural networks with 10 sensory neurons and 3 motor neurons (Fig. 2 Right). Eight sensory neurons encode the activation states of the 8 infrared sensors (which detect obstacles up to a distance

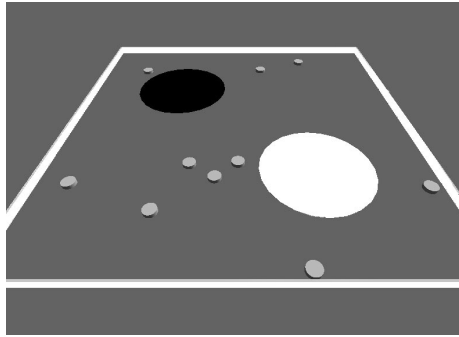


Fig. 1. The environment of the experiment. Black and white circles represent the two foraging areas. Small grey circles represent the 10 robots.

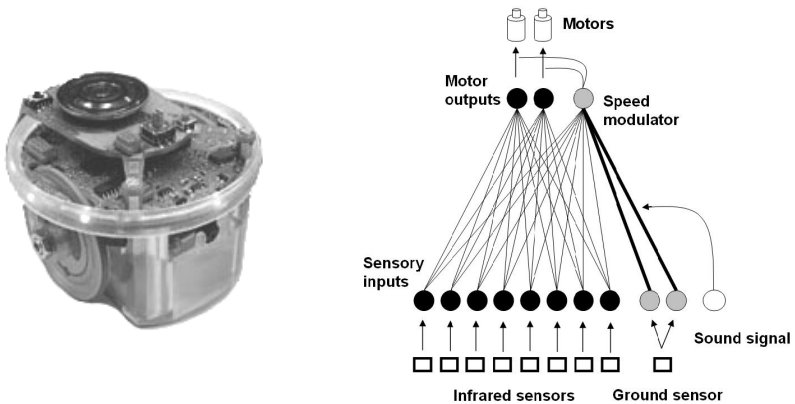


Fig. 2. Left: the e-puck robot – Right: the neural controller. Thick lines represent the weights that are modified in the second phase of the simulation. As it is shown the speed modulator does not operate on the outputs but it acts directly on the motors' activations. The intensity of the sound signal is used to set the learning rate in the social condition.

of ~ 4 cm) and two neurons that measure the activation of the ground sensors that encode the color of the floor below the robot (grey [0 0], white [1 0], and black [0 1]). The first two motor neurons encode the desired speed of the two corresponding wheels normalized between $-\text{MaxSpeed}$ and $+\text{MaxSpeed}$. The third motor neuron instead encodes the maximum speed of both wheels (i.e. the MaxSpeed parameters normalized in the range [0, 10]). The third motor neuron thus acts as a modulator that regulates the speed with which the robot moves or turns at each time step. The 8 sensory neurons that encode the state of the infrared sensors are connected to all the three motor neurons. The two sensory neurons that encode the color of the ground, instead, are connected to the third motor neuron only (i.e. to the motor neuron that regulates the maximum speed of the two wheels).

The robots emit a sound signal through their speakers when they are located in the black or white foraging areas. Moreover, the robots can detect through their microphones the signals produced by other robots located nearby.

2.1 Simulation and Experimental Conditions

The training process is divided in two phases. During the first phase the robots are trained for developing an ability to avoid obstacles (i.e. walls and other robots) and to explore the environment. During the second phase the robots are trained to live as longer as possible in an environment that contains the white and the black target areas. Given that the black area causes a loss of energy, the robots should avoid (i.e. to quickly abandon) it and spend time on the white one. Notice that, in this phase, the behavior of the robots outside the areas has been already determined during the first phase of the training process and cannot be modified, therefore robots can only modify their behavior inside the target areas.

During the first phase, the free parameters that encode the strength of the connections between the 8 infrared sensors and the three motor neurons are selected through an evolutionary method [10]. The robots are placed in an environment without target areas. The free parameters of each robot, encoded with 8 bits, are set randomly and normalized in the range $[-10.0; +10.0]$. The robots are evaluated on the basis of their ability to move and to avoid obstacles in 20 trials after being placed in the environment with randomly selected positions and orientations. The best 20 individuals are allowed to reproduce by generating 5 offspring each. Offspring inherit the same free parameters of their parents but each bit of their free parameter is mutated with a probability of 3%. The evaluation, selection, and reproduction processes are repeated for 50 generations.

During the second phase, the free parameters that encode the strength of the connections between the two ground sensors (that encode the color of the floor) and the third motor neuron (that encodes the maximum speed of the two wheels) are trained through different training procedures corresponding to the five different training conditions described below. In all experimental conditions, individual robots are evaluated for 1 trial lasting 10000 time steps (of 100ms each). At the beginning of the trials the robots are placed in the environment with a randomly selected position and orientation and are provided with an initial energy of 200 units. Every time step spent on the black target causes the loss of 1 unit of energy. When a robot finishes up its energy it “dies” and it is replaced with a new robot placed in the environment in a randomly chosen position and orientation, with full initial energy (the initialization of the free parameters of the new robots depends on the different experimental conditions, see below).

Notice that the robots do not have any direct feedbacks about the “quality” of a target area or about their own energy, i.e. they only have information about the colors of an area, but they do not know the way in which staying in a particular area affects their energy state. In practice, given the exploratory behavior acquired during the first phase, the free parameters that encode the strength of the connection between the two ground sensor neurons and the third motor neuron should be set so to allow the robot to slow down and, eventually, to stop in the white target area and to quickly move out from the black target area.

Individual Learning. The two weights that connect the ground sensor neurons to the third motor are initialized, at the beginning of each trial, with a null value, that is, the robots tend to go away from both the areas. When a robot happens to stay inside a target area, a learning algorithm acts to reinforce, with a positive feedback, the indi-

vidual “preference” for that area, independently from the color of the area and consequently, from the effect on its energy. This positive feedback mechanism is realized by modifying the two free parameters every time an individual is inside one area through the addition of a randomly chosen value in the interval $[-lrate, +lrate]$ to their current value, and by retaining only the variations that lead to an increase of the time spent by the individual in the corresponding target area. This procedure has been tested in two experimental conditions in which the learning rate parameter ($lrate$) has been set to 0.1 and 1.0, respectively. Individual learning has been tested in these two conditions to verify whether the different results obtained in this condition and in the next condition depend on the absolute value of the learning rate or not.

Social Learning. In this condition the free parameters are varied according to the same procedure described above but the learning rate is modulated by the number of other individuals located nearby. More precisely the learning rate is set to a value corresponding to the intensity of the signal detected, which varies linearly in the range $[0.1, 1.0]$ according to the number of individuals located in the same foraging area.

Notice that in both cases (individual and social learning) the learning algorithm is totally “blind” in respect to the adaptiveness of the behaviors. In the former condition, the only way to orient the learning is related to the fact that individuals with adaptive behavior tend to survive and they are not replaced with naïve individuals. In the latter condition, moreover, the robots can exploit an equally blind conformist bias [12], that is, they tend to assume with more probability a behavior if other robots show already this behavior.

Genetic Evolution. In this condition, at the beginning of the experiment, the two free parameters are initialized with randomly selected value in the range $[-10.0; +10.0]$. During the trial, dying robots are replaced with newborn individuals that inherit their free parameters from the robot that currently has the highest energy value and that lives since longer time, with the addition of mutations that are realized by adding to the value that are inherited a randomly selected value in the range $[-2.0, 2.0]$. This value was chosen among a large number of tests as it produces the best performance in this condition (results not shown).

Random Search. This is a control condition identical to the Genetic Evolution condition but in which the free parameters of newborn individuals are assigned randomly in the range $[-10.0; +10.0]$.

For time reasons all the experiments have been carried out using a simulator that carefully reproduces the characteristics of the e-puck robots.

3 Results

By analyzing the obtained results we observed that performance varies significantly in different experimental conditions and that the *Social Learning* condition outperforms all the other conditions (henceforth when we use capital letters and italics we refer to the experimental conditions).

Table 1 shows the average performance obtained in the five experimental conditions at the end of the training process, as well as the average value of some indexes that characterize individuals’ behavior. The Target W and Target B indexes indicate the percentage of the robots of a team located, at the end of each trial, inside the white

and the black target areas, respectively. The mortality index indicates the number of robots that died during the trial, which is inversely correlated with robots' ability to discriminate between the two target areas. The overall performance index is calculated by means of the following equation:

$$W_{perc} + \frac{\left(\frac{M_{MAX} - M_n}{M_{MAX}} \right)}{2} \quad (1)$$

Where W_{perc} is the percentage of the robots of the team located inside the target W at the end of the trial, M_{MAX} is the maximum number of dead robots in all the five experimental conditions and M_n is the number of dead robots in the specific experimental condition. In other words, this index measures robots' ability to minimize the risk of mortality by minimizing the energy lost in the black target area and maximizing the time spent on the white area.

The last two columns of Table 1 show the average value of the two free parameters that regulate the time spent by each robot in the white and black target areas, respectively.

Table 1. Average results and average index values at the end of the training process for the five experimental conditions (see text). Each data represents the average result of 100 trials and the respective standard deviation.

	Target W	Target B	Mortality	Overall Performance	Weight 1	Weight 2
Individual 0.1	0.78 (±0.11)	0.018 (±0.03)	14.7 (±4.35)	0.63 (±0.08)	8.13 (±0.99)	1.95 (±0.56)
Individual 1.0	0.97 (±0.06)	0.003 (±0.01)	12.75 (±6.32)	0.76 (±0.06)	9.71 (±0.58)	0.09 (±0.24)
Social Learning	0.98 (±0.04)	0 (±0)	6 (±0.39)	0.88 (±0.04)	9.86 (±0.39)	1.45 (±0.57)
Genetic Evolution	0.56 (±0.31)	0.037 (±0.06)	26.04 (±10.22)	0.32 (±0.22)	5.63 (±4.58)	0.74 (±4.65)
Random Search	0.37 (±0.14)	0.052 (±0.07)	28.75 (±6.01)	0.18 (±0.10)	1.77 (±1.86)	-1.34 (±1.58)

In particular, by observing the overall performance we can see that the best performances have been obtained by the *Social learning* condition, followed by the *Individual Learning 1.0* and the *Individual Learning 0.1*. These results indicate that, in order to prevent the loss of energy caused by the black area and to minimize the risk of dying, individuals should locate the white target area as fast as possible and as accurately as possible. Moreover, these results indicate how the regulatory effect of social learning on individual learning ensures both speed and accuracy of learning. The fact that *Social Learning* and the two *Individual Learning* conditions significantly outperform the *Genetic Evolution* condition indicates that the latter method is much slower than the formers (for all the Overall Performance results: $p < 0.001$).

By analyzing how Target W (i.e. the percentage of the individuals of a team that are located in the white target) and the mortality indexes vary during the adaptive

process (Figure 3) we can see how the curves referring to the *Genetic Evolution* and *Random Search* conditions substantially overlap during the first phase and start to diverge only after approximately 2200 and 6000 time steps with respect to the two indexes. Moreover we can observe how the index of mortality of the *Social Learning* condition is similar to those of the *Genetic Evolution* and *Random Search* conditions rather than to *Individual Learning* conditions during the first phase but significantly diverges after 2000 time steps.

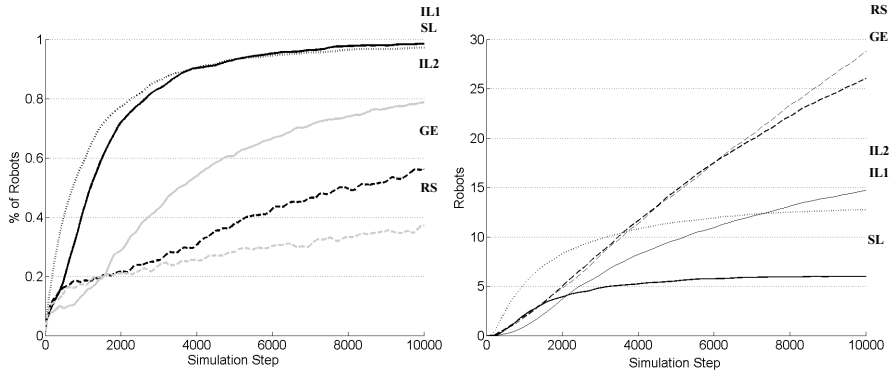


Fig. 3. Left: Percentage of robots located within the white area during the training process. Right: Mortality per trial (i.e. number of robots died during a trial). Both graphs represent the average results of 100 replications for each experimental condition. SL, IL1, IL2, GE, RS indicate the *Social Learning*, *Individual Learning 1.0*, *Individual Learning 0.1*, *Genetic Evolution*, and *Random Search* conditions.

4 Discussion

The results presented in the previous section clearly show that the *Social Learning* condition overcomes the other conditions. The *Individual Learning 1.0* condition leads to similar results with respect to the number of robots located in the white area but it is characterized by a higher mortality rate. Overall the results obtained indicate how the combination of a simple non-directional individual learning process, a simple social mechanism that regulates the strength of the individual learning process, and a selection process that operates at the level of the population (the robots that happen to learn non-adaptive behaviors tend to die and to be replaced by robots with no preference for the white or for the black target) can produce an extremely effective adaptive process.

The individual learning process is non-directional since the mechanism with which it is realized is totally “blind” to whether the variations that are retained are adaptive or not with respect to a single agent per se. Similarly, the mechanism that regulates how social interactions affect individual learning consists of a simple conformist bias [12] which is also “blind” to whether it leads to adaptive or to counter-adaptive variations at the level of a single individual. However, the combination of these two mechanisms together with the selection process ensures that, after an initial phase, the

population of individuals displays an ability to quickly and reliably acquire the required adaptive skill.

In general terms, the mechanisms included in the model described in this paper are also plausible from a biological point of view. Indeed, most animals have the tendency to be neophobic (i.e. reluctant to taste food that they have never eaten before) and to prefer food which they know other conspecifics have eaten [1]. In our model, in fact, at the beginning of the trials that involve learning, the weights that connect the ground sensor neurons to the third motor (the modulator) are initialized with a null value, that is, the robots tend to go away from both the areas. Moreover, as we made clear before, in the *Social Learning* condition the strength of individual learning is modulated by the presence of the other robots in the areas.

The fact that the *Genetic Evolution* condition displays poor performance, on average, can be explained by considering the following three points: (a) the limited size of the population which is composed of only 10 individuals (for an analysis of genetic algorithms in small populations see [10]), (b) the fact that the searching space of good solutions becomes more narrow as the number of robots that show a good solution increases, and (c) the high sensitivity to the initial conditions. The high sensitivity of the results to the initial conditions is demonstrated by the variability of the performance observed in different replications (Figure 4).

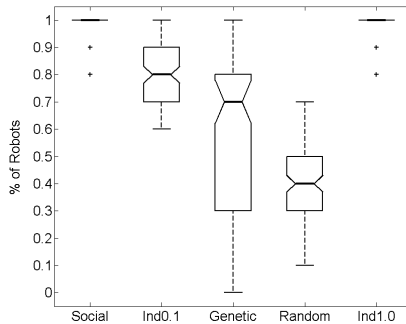


Fig. 4. Distribution of the final percentage of robots in the white area (target W) for 100 trials in the five experimental conditions. Boxes represent the inter-quartile range of the data. The horizontal lines inside the boxes indicate the median values. The horizontal lines outside the boxes indicate the minimum and maximum values. Crosses represent outliers.

The obtained results also indicate that while genetic algorithm applied to small populations might not be effective, the combination of selection at populational level, individual learning, and social learning might produce effective results also in small populations.

The importance of the interaction between the three processes (selection at population level, individual learning, and social learning) can be appreciated by comparing the dynamics in the different experimental conditions. By comparing the mortality rate of the *Individual Learning 0.1* with the mortality rate of the *Social Learning* condition (Figure 3, right) we can see that, during the first 2000 time steps, the latter is substantially higher than the former. This means that, at the beginning of the process, social interactions produce an increase of the number of counter-adaptive behaviors

among the group. In fact, as shown in Figure 5, the percentage of robots located in the black target area at the beginning of the training phase is higher in the case of the *Social Learning* condition than in the case of the *Individual Learning 0.1* condition. This can be explained by considering that social learning act as a conformist bias behavior that tend to strengthen the most common behaviors during this phase, independently from their adaptiveness, since the mechanisms that regulate individual learning and social learning are not necessarily adaptive by themselves. However, the same conformist behavior, combined with the fact that the ratio between adaptive and maladaptive individuals tends to increase thanks to the selection process operating at the level of the population, allows the *Social Learning* conditions to produce better results later on.

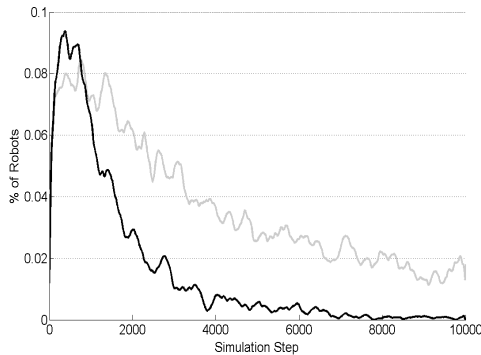


Fig. 5. Percentage of robots inside the black area (target B) through time. Data averaged over 100 trials. The black and gray line indicate the results in the case of the *Social learning* and *Individual Learning 0.1* conditions, respectively.

5 Conclusion and Future Works

We have demonstrated on a simple experimental scenario how social facilitation of individual learning combined with a selection process operating at the population level can be useful in developing adaptive behaviors in small populations of autonomous robots.

In future, we plan to investigate further the scalability of this mechanism in different environmental conditions, e.g., by increasing the number of foraging areas, by increasing (or decreasing) the number of robots, or by increasing the complexity of the perceptual categorization process. We plan to test the ability of the model to cope with changing environmental conditions. Finally, we plan to encode in the genotype of evolving individuals the learning rules that regulate individual and social learning.

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

1. Adessi, E., Visalberghi, E.: How social influences affect food neophobia in captive chimpanzees. A comparative approach. In: Matsuzawa, T., Tomonaga, M., Tanaka, M. (eds.) *Cognitive development in chimpanzees*, pp. 246–264. Springer, Tokio (2006)
2. Dautenhahn, K., Nehaniv, C.L. (eds.): *Imitation in animals and artifacts*. MIT Press, Cambridge (2002)
3. Dugtakin, L.A.: Copying and mate choice. In: Heyes, C.M., Galef Jr., B.G. (eds.) *Social learning in animals: the roots of culture*, pp. 49–64. Academic Press, San Diego (1996)
4. Galef Jr., B.G.: Social enhancement of food preferences in Norway Rats: a brief review. In: Heyes, C.M., Galef Jr., B.G. (eds.) *Social learning in animals: the roots of culture*, pp. 49–64. Academic Press, San Diego (1996)
5. Heyes, C.M., Galef Jr., B.G.: *Social learning in animals: the roots of culture*. Academic Press, San Diego (1996)
6. Nehaniv, C.L., Dautenhahn, K. (eds.): *Imitation and Social Learning in Robots, Human and Animals: Behavioural, Social and Communicative Dimensions*. Cambridge University Press, Human and Animals: Behavioural (2007)
7. Noble, J., Tuci, E., Todd, P.M.: Social learning and information sharing: an evolutionary simulation model of foraging in Norway rats. In: Floreano, D., Mondada, F. (eds.) *ECAL 1999*. LNCS, vol. 1674, pp. 514–523. Springer, Heidelberg (1999)
8. Noble, J., Todd, P.M.: Imitation or something simpler? Modelling simple mechanism for social information processing. In: Dautenhahn, K., Nehaniv, C.L. (eds.) *Imitation in animals and artifacts*, pp. 423–439. MIT Press, Cambridge (2002)
9. Nolfi, S., Floreano, D.: Learning and Evolution. *Autonomous Robotics* 7(1), 89–123 (1999)
10. Nolfi, S., Floreano, D.: *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MIT Press, Cambridge (2002)
11. Reeves, C.R.: Using Genetic Algorithms with Small Population. In: Forrest, S. (ed.) *Proceedings of the Fifth International Conference on Genetic Algorithms*, pp. 92–99. Morgan Kaufmann, San Francisco (1993)
12. Richerson, P.J., Boyd, R.: *Not by genes alone: How culture transformed human evolution*. Chicago University Press, Chicago (2004)
13. Tomasello, M.: *The cultural origins of human cognition*. Harvard University Press, Harvard (2001)
14. Zentall, T.R., Galef Jr., B.G.: *Social learning: psychological and biological perspectives*. Erlbaum, Hillsdale (1988)