We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

## 4,800
Open access books available

## 122,000
International authors and editors

## 135M
Downloads

Our authors are among the

## 154
Countries delivered to

## TOP 1%
most cited scientists

## 12.2%
Contributors from top 500 universities

**BOOK CITATION INDEX**
CLARIVATE ANALYTICS
INDEXED

**WEB OF SCIENCE**™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

## Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

# The Coevolution of Behavior and Motivated Action Selection

Fernando Montes-González, Rodrigo Palacios Leyva,
Fernando Aldana Franco and Victor Cruz Alvarez
*Departamento de Inteligencia Artificial, Universidad Veracruzana*
*México*

## 1. Introduction

The problem of action selection can be identified in ethology as the behavior switching problem. Here, a behavioral repertoire is chosen and then one behavioral module has to be selected until completion or its execution proves ineffective. In behavior based robotics this approach has been widely employed [Arkin (1998)], where an action selection mechanism (ASM) is used to arbitrate between several behavioral modules. The ASM helps in finding an immediate response to the basic question of intelligent systems of what to do next based on its previous sensory perception. However, it is important to notice that the right decision has to be made at the right time [Maes (1990)]. Therefore, some researchers in robotics consider important to develop a model able to decide the right time for making the right decision when solving a particular task. On the other hand, some other researchers are interested in providing a complete solution for the task to be solved. For example Nolfi proposes the evolution of a can-collection task using different motor neurons, and providing some sort of selector, but integrating the solution as a general behavior [Nolfi (1997)]. Potential problems related to the artificial separation of behavioral modules and to the process fusion of the behaviors support the development of general behavior. Also pointed out by the work of [Seth (1998). However, the works of [Kyong Joong & Sung Bae (2001); Yamauchi & Beer (1994)] offer an incremental solution for combining different architectures for solving an entire task. Here, the full integration of evolution with a central action selection model is the main interest.

In this work, the behavioral modules are developed as separate components that can be integrated by the use of a selector. These modules can be implemented as neural networks, programming routines, or a mixture of both. The neural controllers may be optimized by the use of artificial evolution. An important feature of action selection is the emergence of opportunistic behavior [Brooks (1989)] that is not coded in the behavioral modules. We have based our current implementation on a model of central action selection that uses sensor fusion (CASSF) to build a uniform perception of the world in the form of perceptual variables [Montes González, Marín Hernández & Ríos Figueroa (2006)]. However, in this implementation we added the use of two simulated motivations which correspond to fear and hunger. Also, the model is able to switch behaviors in normal and lesioned conditions (not explored in this paper). At this point we have decided to use staged evolution of behavior and then coevolution to adjust the weights of the neural controller for the selection mechanism. Hence, the evolutionary robotics approach was used in the design of three

exploration behavioral patterns: one for finding walls, another to locate corners, and other for locating cylinders. Then, due to its sequential nature the design of the collection and deposit behavioral patterns was implemented as algorithmic routines.

The optimization of action selection and behavioral modules present differences with a standard hand-coded implementation. For instance, in a previous work hand-coded selection produced regular patterns of behavior [Montes González, Santos Reyes & Ríos Figueroa (2006)]. Non-regular patterns of behavior were produced using coevolution to optimize both behavior and selection [Montes González (2007)]. However, instability in the system was avoided by the execution of selection in a non-motivated environment. Here, we show how motivated behavior can be optimized by the use artificial evolution.

## 2. Action selection and evolutionary robotics

The action selection problem is related to granting control of available actuators to behavioral modules while the stimuli that activated them is valid or wears out. In the vertebrate brain, at specific loci, specialized centers of selection can be identified. One of them is the basal ganglia, and recent works support the idea of these nuclei playing an important role in action selection [Prescott et al. (2006)]. The basal ganglia act as a relay station in the planning and the execution of movements (behavior); hence gathering information from the cortex and motor cortex. The basal ganglia are able to mediate cognitive and muscular processes. Not only serves the basal ganglia as an important center of action selection also in cooperation with the cerebellum and the sensory cerebrum; all of them are able to veto muscular contraction by denying the motor areas sufficient activation. In turn, these individual motor elements form more complex patterns, which can be thought as essential elements in the development of intelligence [Bares & Rektor (2001)]. The development of intrinsic basal ganglia circuitry with evolvable behavioral modules has already been explored in [Montes González (2007)].

The use of artificial evolution in robotics requires an implementation based on Evolutionary Algorithms. Outstanding representatives of the Evolutionary Algorithms are Genetic Algorithms (GAs), Evolution Strategies, Genetic and Evolutionary Programming and Coevolution [Bäck (1996); Santos & Duro (2005)]. However, in this study the use of GAs [Holland (1975)] and Coevolution [Ebner (2006)] were preferred. In the existing literature examples of evolutionary algorithms are provided for the modeling of elementary behavior controllers [Floreano & Mondana (1996); Seth (1998); Yamauchi & Beer (1994)]. The use of neural networks with genetic algorithms is a common approach for providing a solution to the modeling of robot behavior [Nolfi & Floreano (2000)]. The use of this approach requires the right choice of a topology and the neural weights to control the robot. Evolutionary Robotics (ER) employs several iterations to refine random individuals that are neural controllers for modelling autonomous robot tasks that are solved in semi-controlled environments. In turn, individuals form a changing population that survives across generations according to their adaptability (fitness) to solve a particular task. Therefore a population of these controllers forms the elementary building blocks [Goldberg (1989)] that will serve as the necessary leverage to increment the overall fitness population. Parameter interactions in the implementation of a robot controller epistatically increment the chances to find near-optimal solutions. Most of the times a single objective evaluation function is employed in ER due to the fact that the resultant robot behavior comes from a dynamic system made with the robot and its environment [Doncieux & Mouret (2010)]. However, in evolutionary computation is possible to simultaneously optimize several objectives without aggregating them as a single monolithic fitness function [Deb (2001)]. Recent research relies on the concept of domination

and generates the so-called Pareto front to bootstrapping a light-seeking robot behavior [Mouret & Doncieux (2009)].

In this paper we use a single objective function to shape a global behavior to be achieved. As a consequence the use of GAs is sufficient to model complex robot behavior that provides a solid approach for starting to work on ER. In this approach is important to notice that tuning the obstacle avoidance behavior, which is shaped and nearly optimized by the use of the GA, ultimately depends on the right choice of the fitness function [Reynolds (1994)]. Furthermore, evolution not only optimizes specific features of individual neural controllers; also pairs or groups of individuals can be evolved together. The evolution of one individual can be affected by the change of other related individuals in the group [Lapchin & Guillemaud (2005)]. The latter approach has been identified, as its biological counterpart, as coevolution that can be cooperative or competitive. A cooperative strategy can be developed to achieve a common task (e.g. pushing objects, solving a common task), whereas in a competitive strategy individuals have to struggle to assimilate some scarce resources (e.g. prey and predator, securing of food stashes). In biology diffuse coevolution has been referred to species evolving in response to a number of other species, which in turn are also evolving in response to a set of species [Ridley (2003)].

Cooperative individuals not only require a society interaction, but the existence of an internal mechanism (e.g. among other CASSF and the computational model of the basal ganglia) that is able to mediate between various sensory processes. Therefore, individuals need to build up unified internal perceptions based on their available sensory capabilities in order to produce specialized behavior. As a consequence sensory processes need to be augmented when possible. The work of [Montes González et al. (2008)] shows how non-standard avoidance can be achieved by extending infrared sensory information through an evolutionary refinement.

## 3. The development of behavior using evolution

The optimization of behavior was carried out using an incremental approach following the evolutionary robotics methodology [Nolfi & Floreano (2000); Santos & Duro (2005)]. Initially modules were evolved in the Webots robot simulator [Cyberbotics (2011)], and later on the modules were further evolved in the real Khepera robot [Mondana et al. (1993)]. The Khepera has been equipped with a ring of eight infrared sensors distributed around the body of the robot and two DC motors that control the movement of the wheels. A foraging task was set in a square walled arena where the robot has to collect simulated "food" in the form of wooden cylinders. Thus, we employ the gripper turret attachment for the Khepera (see Figure 1), this small arm has two degrees of freedom with encoders for determining the arm position and two sensors in the gripper claw for detecting the presence and the resistivity of a collected item. Behavior can be identified as belonging to two different kinds, some related to travelling the arena and the other related to handling objects with the gripper. Hence, exploration behavior implemented as neural behavior corresponds to *cylinder-seek* which locates and positions the robot body in front of a cylinder in order to collect an object; *wall-seek* travels the arena searching for the closest wall, and then *corner-seek* runs parallel to a wall until the robot finds a corner. In order to complete the behavioral repertoire collection and release of cylinders is carried out by *cylinder-pickup* that moves the robot backwards to safely lower the robot arm and then pick up a cylinder; finally *cylinder-deposit* lowers the robot arm, opens the gripper and returns the arm to an upper position.
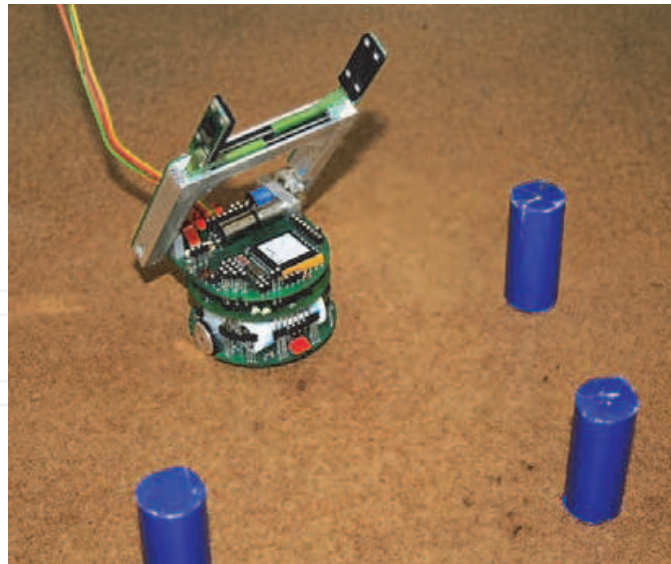
Fig. 1. The robot is shown in a walled arena where has to collect wooden cylinders. Thus the robot has been fitted with a gripper turret attached to the Khepera robot.

### 3.1 Neural behavior

In our experiments we use the infrared sensors of the Khepera robot to transverse the arena. Thus, we chose to implement exploration behavior as neural controllers. At an initial stage we individually evolve the next behavioral modules *wall-seek*, *corner-seek* and *cylinder-seek*. These modules use a fully connected feedforward multilayer-perceptron neural network with no recurrent connections. The topology of the neural network is six neurons in the input layer, five neurons in the hidden layer, and two in the output layer. The sigmoid transfer function is used at the hidden and the output neurons. The infrared output from the Khepera, ranging from 0 to 1023 from the six frontmost sensors forms the input to the neural neural network. Then, the output of the neural network is scaled to the $\pm 20$ values required for driving the DC motors at full speed. Next, a genetic algorithm with *selection*, *crossover* and *mutation* operators was applied to the neural network and the desired behavior for each individual module was shaped using different fitness functions (Eqs. 1, 2 and 3).

The weights of each neural network are directly encoded into a vector **w** of 40 elements, the weights were initialized with random values ranging from $-1 < \mathbf{w}_i < 1$ for all elements. Thus, a single vector representation is used to define each of the individuals in the population. The initial population, $G_0$, consists of $n=100$ neural controllers. Selection is made using *elitism* to replicate the two best individuals from one generation to the next. Next, a *tournament* allows random parents to be chosen from $(n/2)$-1 competitions. The most fitted parents are bred in pairs with a random crossover point generated with a probability of 0.5. Each individual in the new population is then affected with a mutation probability of 0.01. The fitness is scored by running each individual in the simulator for about 30 seconds. The initial location and orientation of the individuals are randomized across trials.

The behavior for locating a wall (*wall-seek*) can be seen as a form of obstacle-avoidance due to the fact that the arena has to be explored avoiding cylinders. Then, the behavior is completed when the robot is in front of a wall. The fitness formula for this behavioral module was

$$f_{c1} = \sum_{i=0}^{4000} abs(ls_i) * (1 - \sqrt{ds_i}) * (1 - max\_ir_i) \tag{1}$$

where for iteration *i: ls* is the linear speed in both wheels (the absolute value of the sum of the left and right speeds), ds is the differential speed on both wheels (a measurement of the angular speed), and *max_ir* is the highest infrared normalized-value. A formula such as this favors the evolution of fast individuals that run in a straight line while avoiding obstacles.

The behavioral module for running parallel to a wall makes the robot move in a straight line aside a located wall; though any obstacles blocking a straight path to the nearest corner have to be first avoided. The module is stopped when a corner is detected. The fitness formula employed for the behavior *corner_seek* was

$$f_{c2} = f_{c1} * (tgh)^2 \qquad (2)$$

This formula employs a thigmotaxis factor *(tgh)*, which accounts for the tendency to remain next to walls and is calculated as the fraction of the test period for which an individual is close to any of the walls in the arena. This formula therefore evolves individuals that avoid obstacles while traveling parallel to the arena walls.

The *cylinder-seek* behavior explores the arena avoiding walls until it locates a cylinder set in the middle of the arena. If a cylinder is located (detected by the two frontmost pair of infrared sensors), then the robot stops to let the gripping-behavior handle cylinder collection. The formula for a behavioral module such as this was

$$f_{c3} = f_{c1} + K_1 * cnear + K_2 * cfront \qquad (3)$$

In this formula avoidance is displayed for travelling the arena. The constants $K_1$ and $K_2$, with $K_1 < K_2$, are employed for rewarding the robot when a cylinder is detected around the ring of infrared sensors assuming that a cylinder is near *(cnear)*. However, the robot is most rewarded when aligns its frontal part with a nearby cylinder *(cfront)*.

### 3.2 Fixed sequences of behavior

The previous behavioral modules can be considered as timed sequences of action triggered by an initial sensory stimulus. However, behavior related to handling the gripper should be modeled as a sequence of particular actions executed always in the same order and with the same duration. Thus, behavior modeled in this way can be thought as fixed action patterns [Mcfarland (1993)]. For instance, *cylinder-pickup* requires the gripper claw to be opened, and then the robot move backwards to create free space in front of the body, the gripper closed, and the arm moved back into the upright position. *Cylinder-deposit* requires a fixed sequence of lowering the arm, opening the gripper, and then raising the arm. Therefore, these two behavioral modules were implemented as algorithmic routines following the aforementioned action sequences.

## 4. Central action selection

In writing different models have been proposed to design systems, which are able to exhibit a variety of behavior and to arbitrate between them [e.g. Arkin (1998); Brooks (1986); Maes (1989)]. Nevertheless, these models based on explicit design do not seem to be scalable enough for developing systems capable of displaying a large variety of behavioral patterns that cope with task/environmental variations. In previous research we have proved that a computational model of the intrinsic circuitry of the vertebrate basal ganglia [Prescott et al. (1999)] produces action selection when embedded in a robot control system [Montes González et al. (2000); Prescott et al. (2006)]. The motivated robot basal ganglia model has been set in a

similar environment to the one described using hand-coded [Prescott et al. (2006)] and evolved behavioral patterns [Montes González (2007)]. The importance of the basal ganglia in natural action selection becomes evident when we observe that these nuclei are an archaic feature common to all vertebrate animals [Prescott et al. (1999)]. However, we have also worked in an alternative selection model named CASSF [Montes González & Marín Hernández (2004)] that shares common features with the robot basal ganglia model. Both are centralized and both produce motor selection based on building perceptual information from raw sensory input. One of the main features of CASSF (Central Action Model with Sensor Fusion) is that is modular and able to cope with the variations of a dynamic environment. However, in this study we have extended CASSF to include internal motivations for the calculation of motor selection (Figure 2). In addition, CASSF is an effective action selection mechanism [Montes González, Santos Reyes & Ríos Figueroa (2006)] that is centralized and presents sufficient persistence to complete a task. The implementation of tasks such as foraging can be carried out in CASSF by determining a set of behavioral patterns that can be integrated in time to complete such a task. The adjustment of selection parameters and behavior has been optimized by coevolution in CASSF as described in [Montes González (2007)].

The foraging activity for our behavioral setup has been modeled loosely based on observations of hungry rats placed in a box containing a central small dish of food. These animals, even when deprived from food for twenty four hours, will be fearful and exhibit preference of staying next to walls and corners. Later on, they will go across the arena to collect food from the dish that is then consumed in a corner. The urgency to be selected (salience) of each of the behavioral modules is tuned to provide appropriate behavioral selections that simulate the avoidance-related and food-acquisition-related behavior observed in these animals. Therefore, the salience for each module depends on the values of a number of *extrinsic* and *intrinsic* variables. Extrinsic values are calculated as bi-polar perceptual variables calculated from robot raw sensory information. These perceptual variables are labeled as *wall_detector*, *gripper_sensor*, *cylinder_detector* and *corner_detector*. Additionally, these perceptual variables are also sent to each of the behavioral modules. The information of the sensors is updated at every step of the simulation and the perceptual variables are recalculated depending on the presence (+1) or absence (-1) of the relevant target feature (e.g. a cylinder, a wall, a corner, or an object in the gripper).

Intrinsic variables are produced by motivational modules and are functions of recent experience and internal state. In our experiments these roughly model "fear" (initially high and reduced when exploring the arena) and "hunger" (increases with time and reduced when "food" is deposited outside the arena). Therefore, the value for each of the simulated motivations is a single scalar value in the range (0-1) that can be either increased or decreased over time. Hunger is also reduced by a fixed amount when a cylinder is deposited in a corner of the arena. On the other hand, behavioral modules are also able to generate an intrinsic variable (a "busy signal") that facilitates its own selection during critical phases of activity. The value of the busy signal is a binary value that is on when a critical period of activity has been reached. As a result the salience is calculated from the relevant information for each behavioral module composed by perceptual variables (bi-polar), its own busy-signal (binary) and extrinsic motivations (scalar values). These signals constitute the input vector for the selection network and activation is computed at every step of the main loop. Thus, CASSF runs within a main loop in which sensor readings are updated and motor commands are sent. At each time-step, salience is calculated and the competition between behavioral components is resolved in a *winner-takes-all* manner.
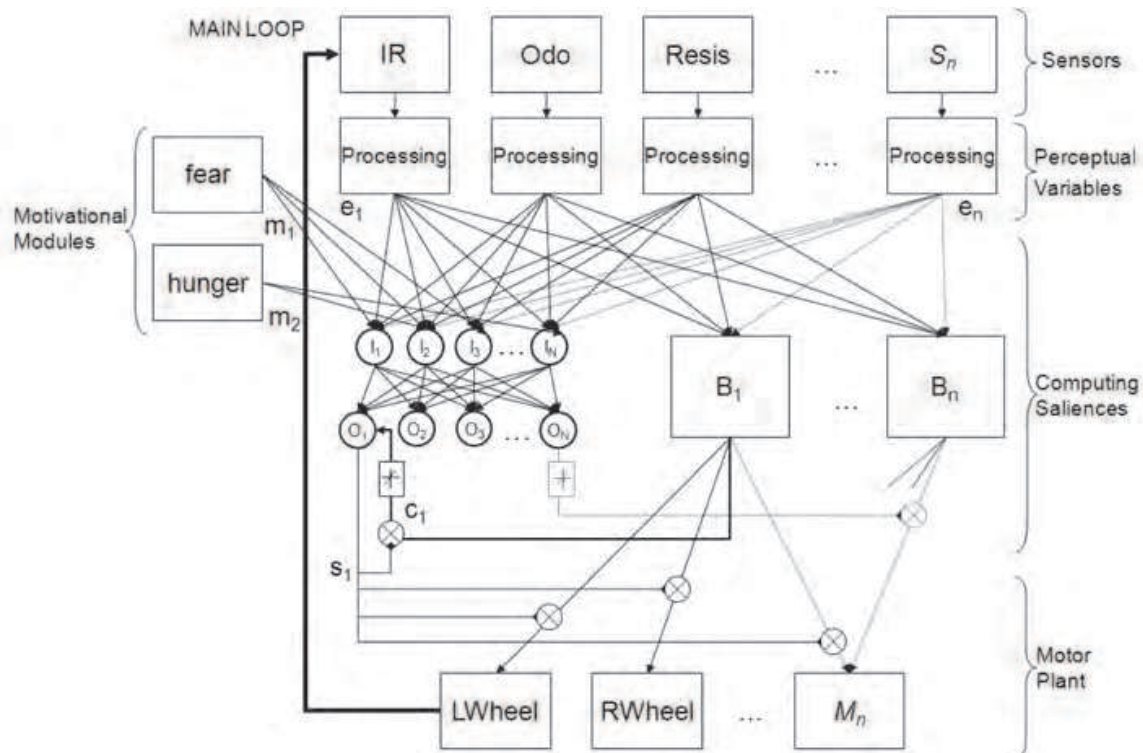
Fig. 2. In the CASFF model, perceptual variables ($e_i$) form the input to the decision neural network. The output of the selected behavior with the highest salience ($s_i$) is gated to the motors of the Khepera. The busy-status signal ($c_1$) from behavior $B_1$ to the output neuron $O_1$ should be taken into account. The behavioral repertoire ($B_1$ - $B_n$) is extended by preserving similar connections for each of the additional behavioral modules. Motivations ($m_i$) are added as inputs ($I_i$) to the decision network.

The calculation of salience is carried out using a decision neural network with an input layer and an output layer both with five neurons (Figure 2). The output of the neurons make use of the identity function to preserve the original calculation of the salience. The raw sensory information from the Khepera is fed into the neural network in the form of perceptual variables. Next, the input neurons distribute the variables to the output neurons. The behavior that is selected sends a busy signal to the output neurons when its salience is above the salience of the other behaviors. A selected behavior sends a copy of this busy signal to the five output neurons. In turn the five behavioral modules can be selected, thus each of the five modules add five more inputs to the output neurons. Additionally, two more inputs are added from the two simulated motivations. Hence, salience output is calculated from the five salience signals of the behavioral modules plus four perceptual variables plus two simulated motivations.

In this paper artificial coevolution is employed to adjust the weights of the decision network. The exploration behavioral modules were evolved as an initial stage in the evolution. A second stage consisted in evolving both the decision network weights and the behavioral networks. The decision network weights depend on the input of the context vector ($e_i$) which is formed by the perceptual variables. In turn, the variables wall_detector ($e_w$), gripper_sensor ($e_g$), cylinder_detector ($e_c$), and corner_detector ($e_r$) are encoded from readings of different sensors. These perceptual variables form the context vector, which is constructed as follows ($e = [e_w, e_g, e_c, e_r]$, $e_w, e_g, e_c, e_r \in \{1, -1\}$). Then, five different behavioral modules return a

current busy-status ($c_i$) indicating that ongoing activities should not be interrupted. The current busy-status vector is $\mathbf{c} = [c_s, c_p, c_w, c_k, c_d]$, $c_s, c_p, c_w, c_k, c_d \in \{1, 0\}$, for cylinder-seek, cylinder-pickup, wall-seek, corner-seek, and cylinder-deposit respectively. The motivational vector is composed by $\mathbf{m} = [m_f, m_h]$, $-1 < \mathbf{m}_i < 1$. The salience ($\mathbf{s}_i$) or urgency is calculated from the input of the decision network $\mathbf{I}_i$ which in turn modifies the output $\mathbf{O}_i$ of the behavioral modules by allowing the most salient to win the competition. Thus, selection is evolved for five behavioral modules with a context vector composed of eleven elements making a selection vector $\mathbf{ch}_s$ of 55 weights with initial random values of $\mathbf{ch}_s$, $-K_w < \mathbf{ch}_s < K_w$ with $K_w = 0.75$. Therefore, the full chromosome to be coevolved is formed as an array of 55 weights of the decision network plus 40 weights of cylinder-seek plus 40 weights of wall-seek plus 40 weights of corner-seek giving a total of 175 weights.

Evolution was carried out as previously described in section 3.1. The fitness formula for the evolution of the weights of the decision network was

$$f_{c4} = (K_1 * avfactor) + (K_2 * cwfactor) + (K_3 * pkfactor) + (K_4 * dpfactor) \tag{4}$$

The evolution of the weights of the selection network was nearly optimized using in the fitness formula $f_{c4}$ the constants $K_1, K_2, K_3$ and $K_4$ with $K_1 < K_2 < K_3 < K_4$ for the selection of those individuals that move in a traight line while avoiding obstacles (*avfactor*), also those individuals that locate corners and walls in the arena (*cwfactor*). Moreover, the fitness formula also rewards cylinder-collection inside the arena *(pkfactor)*, and their release near the outside walls *(dpfactor)*. The average fitness of a population, for over 100 generations, and its maximum individual fitness are shown in Figure 3.
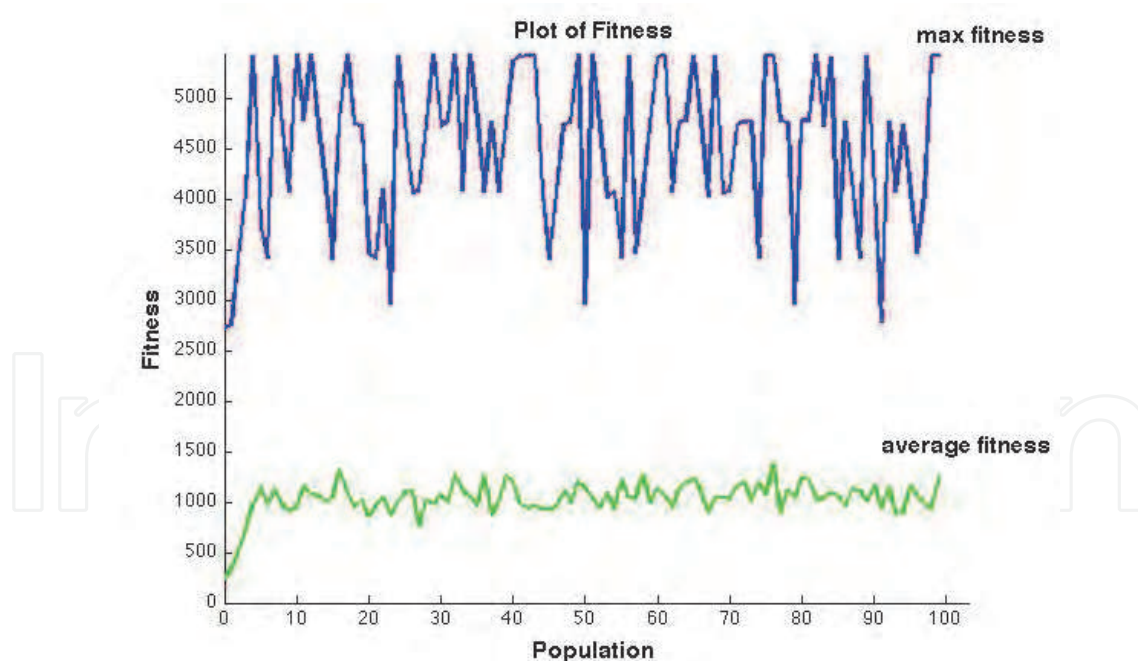


Fig. 3. Fitness is plotted across 100 generations. For each generation the highest fitness of one individual was obtained from the averaged fitness of five trials under similar conditions. The maximum fitness of all individuals was averaged as a measure of the population fitness. Individuals are more rewarded if they avoid obstacles, collect cylinders, and deposit cylinders close to corners. The evolution is stopped after fitness stabilizes over a value around 4500.

## 5. Experiments and results

The foraging task was set in an arena with four cylinders as simulated food (see Figure 4). The use of selection with evolved parameters is shown in Figure 5. On the other hand, the coevolution of behavior and selection parameters is presented in Figure 6. It is important to notice that coevolution tends to optimize behavior and their selection in time and in the physical environment. Therefore, it is possible to observe that in attempt to maximize fitness; evolution may disrupt the order of selection as anticipated by the human designer. The latter, has been pointed out in [Montes González, Santos Reyes & Ríos Figueroa (2006)] where five behavioral modules were setup for coevolution. The presence of redundant behavior leads to an evolutionary race, because interference from other behavior may affect the development and selection of a particular behavior. A regular behavioral pattern may be easily observed if we were using hand-coded selection whereas for evolved and coevolved selection the use of corner-seek is avoided in order to optimize timed selection. The selection of behavior patterns is also summarized as elementary statistics in tables 1 and 2 for evolved selection and coevolved selection/behavior. In this paper a behavior is considered as the joint product of the robot and its internal status, environment, and observer. Hence, a regular grasping-depositing pattern in the foraging task should be the result of the selection of the behavioral modules: cylinder-seek, cylinder-pickup, wall-seek, corner-seek, and cylinder-deposit in that order. Collection patterns can be disrupted if for example the cylinder slips from the gripper or a corner is immediately found. Another cause for a change in a regular pattern occurs after long search periods when a cylinder is not promptly located. The use of motivations is also another cause for the interruption in the collection of a cylinder. For instance travelling for long time increases the value of hunger up to its maximum value, which makes locating a cylinder erratic and increasing periods of exploration. Finally, the fitness of the agent solving the foraging task is an additional factor that alters the order in the selection of behavior.

In Figure 7 we present a comparison of evolved-selection and coevolved-selection/behavior. In this comparison we notice that even though individuals present different fitness values, they are able to complete the task of collecting from one to four cylinders. Fitness is increased after long search periods, and those individuals able to complete the same task with lower fitness are because they have to travel less to locate cylinders. Consequently, "lucky" collectors, by chance, travel less earning fewer rewards; whereas "unlucky" collectors travel more earning additional rewards. As a result, in the same figure we observe an improvement on the fitness and the collection of cylinders of coevolved selection in comparison to evolved selection and behavior.

## 6. Discussion and future work

For our discussion it is important to remember that there is evidence of central selection in the vertebrate brain [Prescott et al. (1999)], particularly at the basal ganglia buried under the cortex. It is important to notice that these structures receive information from several different regions of the cerebral cortex. We have based the development of motivated CASSF on that of the robot basal ganglia [Prescott et al. (2002)]. Therefore, we have implemented an action selection model that makes use of artificial evolution for optimizing both neural behavior and the decision network (ruling out the evolution of sequential behavior). In our selection model we have build an intrinsic perception of the world based on raw sensory information to provide pre-processed information to the decision network in order to produce a unified perception of the "extrinsic" world. Additionally, we have let intrinsic variables such as
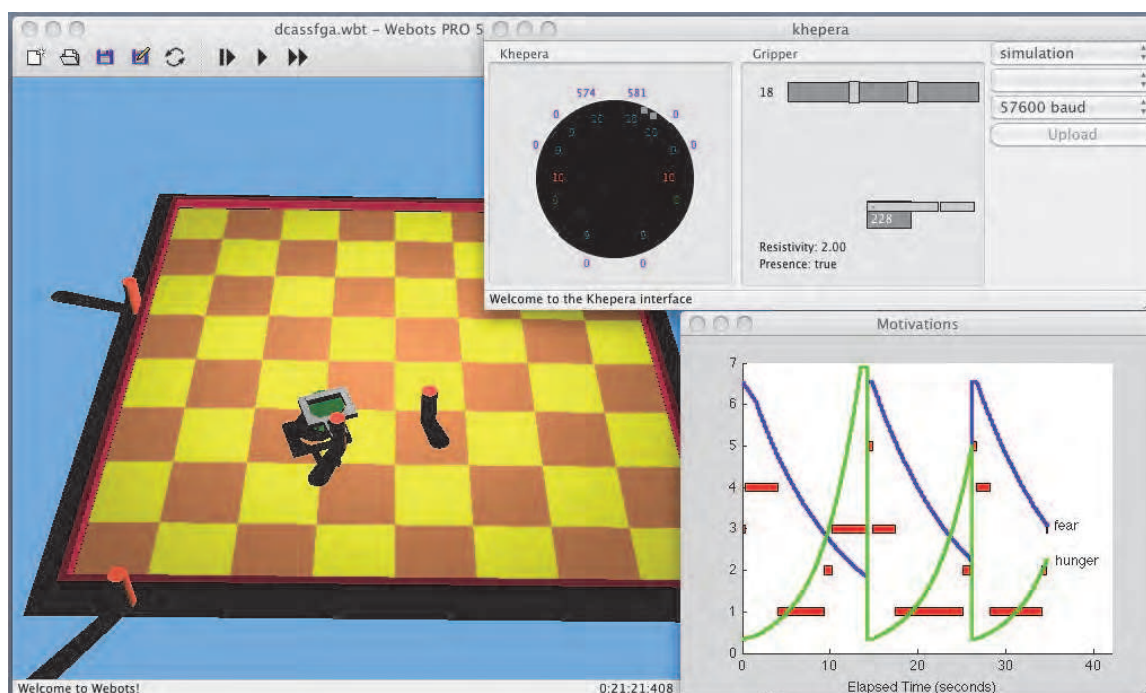
Fig. 4. The selection of behavioral modules and their related motivations are shown in the image. The robot simulator window and the window for the gripper and infrared sensors are also displayed. In the motivations window, the blue line corresponds to "fear" and the green line represents "hunger". Final behavior is then transferred to the real robot for further optimization.

simulated fear and hunger to affect the results of selection. Therefore, selection arbitrates among competing behavioral modules to allow the execution of behavior in response to a specific configuration of the world and the internal status of the animal robot.

| Behavioral Modules | Freq | Latency | TotDur | TotDur% | Mean | StdDev | StdErr | MinDur | MaxDur |
|---|---|---|---|---|---|---|---|---|---|
| none | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| cylinder-seek | 22.00 | 0.00 | 61.88 | 72.80 | 2.81 | 5.46 | 1.16 | 0.00 | 22.03 |
| cylinder-pickup | 13.00 | 7.50 | 13.11 | 15.42 | 1.01 | 1.09 | 0.30 | 0.00 | 4.06 |
| wall-seek | 11.00 | 0.00 | 5.63 | 6.62 | 0.51 | 0.33 | 0.10 | 0.16 | 1.41 |
| corner-seek | 0.00 | 85.00 | 0.00 | 0.00 | NaN | NaN | NaN | 0.00 | 0.00 |
| cylinder-deposit | 4.00 | 14.53 | 4.38 | 5.15 | 1.10 | 0.01 | 0.00 | 1.09 | 1.10 |
| Total | 51.00 | 0.00 | 85.00 | 100.00 | 1.67 | 3.73 | 0.52 | 0.00 | 22.03 |

Table 1. Evolved selection (Figure 5) is summarized in this table. Labels in the table are as follows: *Freq* shows the frequency in the selection of a module; *Latency* represents the time when the module was initially selected; the total duration of the module is indicated by *Totdur* and its percentage by *TotDur%;Mean*, Standard Deviation *(StdDev)* and Standard Error *(StdErr)* are some simple statistics; *MinDur* represents the minimal time the module was selected and *MaxDur* the maximal time for the selection of the module. Here we observe that cylinder deposit was activated only four times corresponding to the collection-release of the four cylinders set in the arena.

Evolutionary robotics employs a quasi-optimal approach to develop autonomous controllers for different kinds of robots and groups of individuals which can be evolved together. Following this approach a change in the evolution of one individual can be affected by the change of other related individuals in the group. The latter approach has been identified, as its biological counterpart, as coevolution that can be cooperative or competitive. A cooperative
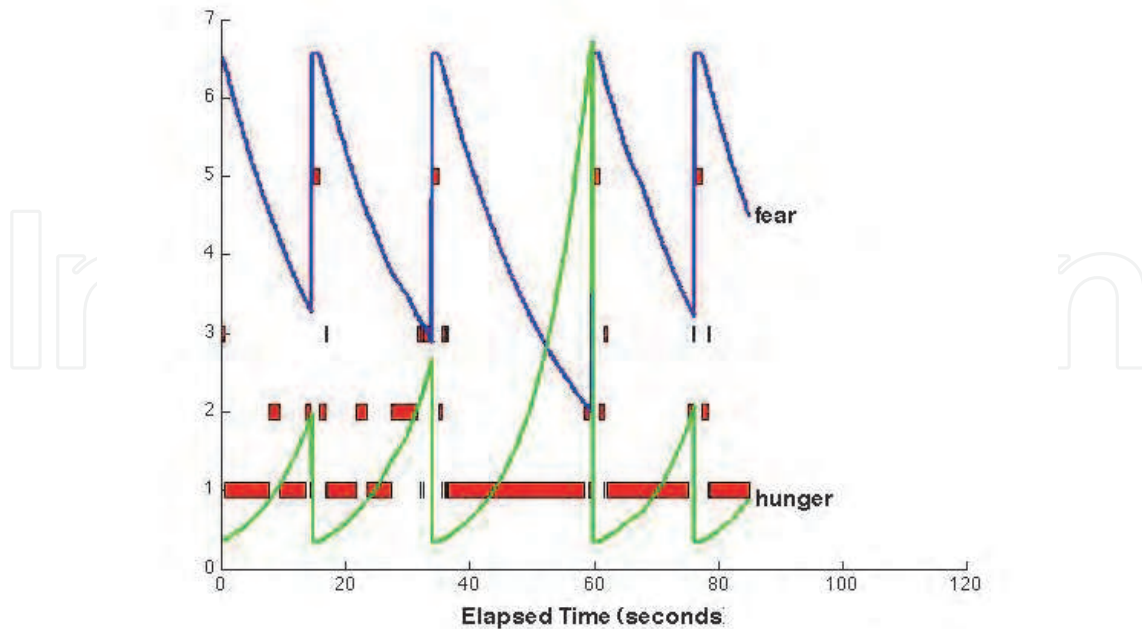
Fig. 5. Ethogram for a run of typical evolved selection; modules are numbered as 1-cylinder-seek, 2-cylinder-pickup, 3-wall-seek, 4-corner-seek, 5-cylinder-deposit and 6-no action selected. We notice that although cylinder-pickup was activated thirteen times, only four cylinders were collected. The latter is due to motivations causing variations in the internal variables thus affecting salience calculation causing the robot to rapidly select and deselect cylinder-pickup several times before actually grasping a cylinder. This individual presents a semi-regular grasping-depositing pattern (behavioral pattern selection of 1-2-3-5).

strategy can be developed to achieve a common task (e.g. pushing objects, solving a common task), whereas in a competitive strategy individuals have to struggle to assimilate some scarce resources (e.g. prey and predator, securing of food stashes). In biology diffuse coevolution has been referred to species evolving in response to a number of other species, which in turn are also evolving in response to a set of species. Consequently, the identification of collective intelligence is more evident when coevolving cooperative groups.

| Behavioral Modules | Freq | Latency | TotDur | TotDur% | Mean | StdDev | StdErr | MinDur | MaxDur |
|---|---|---|---|---|---|---|---|---|---|
| none | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| cylinder-seek | 113.00 | 0.46 | 54.66 | 60.42 | 0.48 | 2.12 | 0.20 | 0.00 | 15.47 |
| cylinder-pickup | 67.00 | 3.28 | 14.84 | 16.41 | 0.22 | 0.66 | 0.08 | 0.00 | 3.91 |
| wall-seek | 55.00 | 0.00 | 16.43 | 18.16 | 0.30 | 0.20 | 0.03 | 0.00 | 1.10 |
| corner-seek | 0.00 | 90.46 | 0.00 | 0.00 | NaN | NaN | NaN | 0.00 | 0.00 |
| cylinder-deposit | 4.00 | 8.12 | 4.53 | 5.01 | 1.13 | 0.20 | 0.10 | 0.94 | 1.41 |
| Total | 240.00 | 0.00 | 90.46 | 100.00 | 0.38 | 1.51 | 0.10 | 0.00 | 15.47 |

Table 2. Elementary statistics for evolved selection; we notice several activations of cylinder-pickup. However, only four cylinders were collected and released outside the arena (notice that cylinder-seek is selected only four times). A proper release of a grasped cylinder corresponds in the image to a decay in the fear motivation (around seconds 8, 29, 64 and 76 in Figure 6). Additionally, evolution avoids the use of corner-seek and that is the reason for its latency to be shown as the total elapsed time even though selection never occurred.

The development of collective behavior in simulation can be achieved by simple scalable control systems as a form of decentralized control. Therefore, the work of [Kube & Zhang (1993)] exhibited group behavior without the use of explicit communication. A more recent
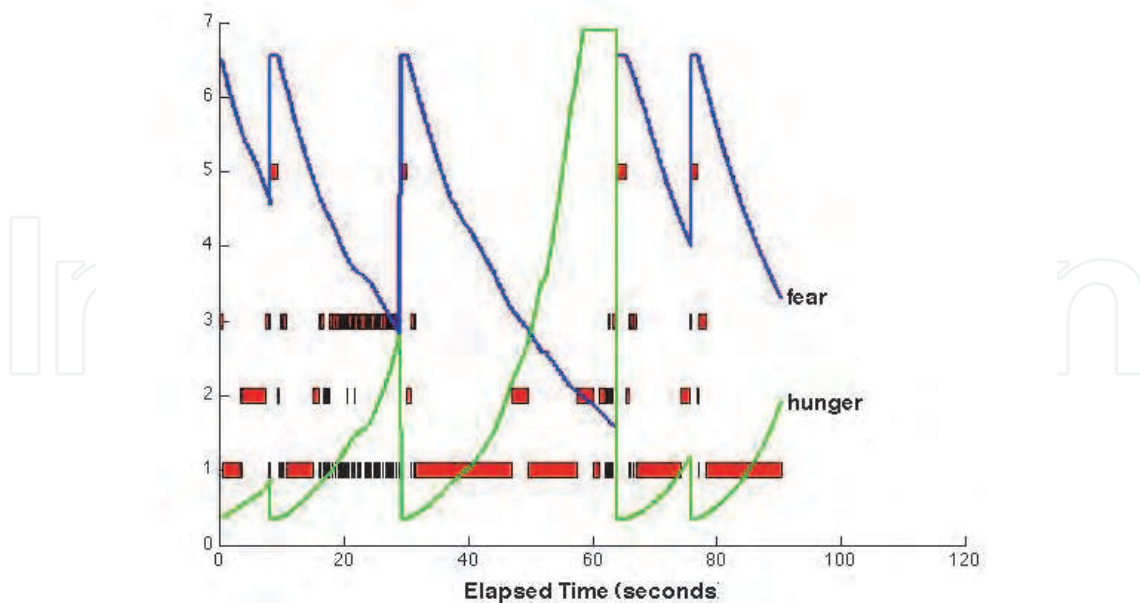
Fig. 6. A typical run for coevolved selection; the behavioral modules are numbered as 1-cylinder-seek, 2-cylinder-pickup, 3-wall-seek, 4-corner-seek, 5-cylinder-deposit and 6-no action selected. A standard grasping-depositing pattern (1-2-3-4-5) is also not observed here. Because of opportunism presented in the collection of cylinders, artificial evolution optimizes the use of corner-seek that is never selected. Instead wall-seek is employed, after collection, for taking the cylinder to the nearest wall where the cylinder is released. It is important to notice that cylinder collection is a bit more erratic but in the end the value of fitness is greater compared to evolved selection. As a result, a semi-regular behavioral pattern selection of 1-2-3-5 is observed in the graph.

approach by Marocco et al. [Marocco & Nolfi (2007)] set a collective task where a group of four robots developed the ability to reach two different target areas and to group in pairs at each different location. Elementary communication skills were evolved alongside the abilities to reach target areas. In this way, coordination is achieved through stigmergy; nevertheless more elaborated communication can be implemented [Mitri & Vogt (2006)]. In general, the individual ability to communicate may be one of the requirements for the development of collective intelligence. The evolutionary approach has been sufficient to solve problems were cooperation and communication skills are necessary for solving a particular task; in contrast communication arises as a deceptive feature within a competitive scenario. In this scenario a common setting is that of the prey and the predator where both individuals are competing for scoring points either for escaping or capturing each other. The experiment can be expanded to add more preys and predators.

The presence of collective intelligence is found in different areas ranging from biology, sociology, and business to computer science. However, a common definition looks for the identification of certain "intelligence" derived from the efforts of joined entities. The aforementioned entities range from bacteria, animals, human to computer processes. Moreover, the same definition can be applied, in a broad sense, to action selection and evolutionary robotics. The emergence of collective intelligence based on the behavior-based approach requires stepping out from the modeling of selfish solitary individuals to social organisms. Therefore, we need to group our robots and expose them to numerous interactions

to assure complex performances at the level of the group. Here, we have identified some common elements in collective robotics: cooperation, intelligence, communication skills, and the integration of sensory information with action selection. All of those accomplished with the use of the evolutionary robotics approach. As a consequence, we believe that we may contribute to the development of robotic collective intelligence by way of social experiments using the artificial evolutionary method.
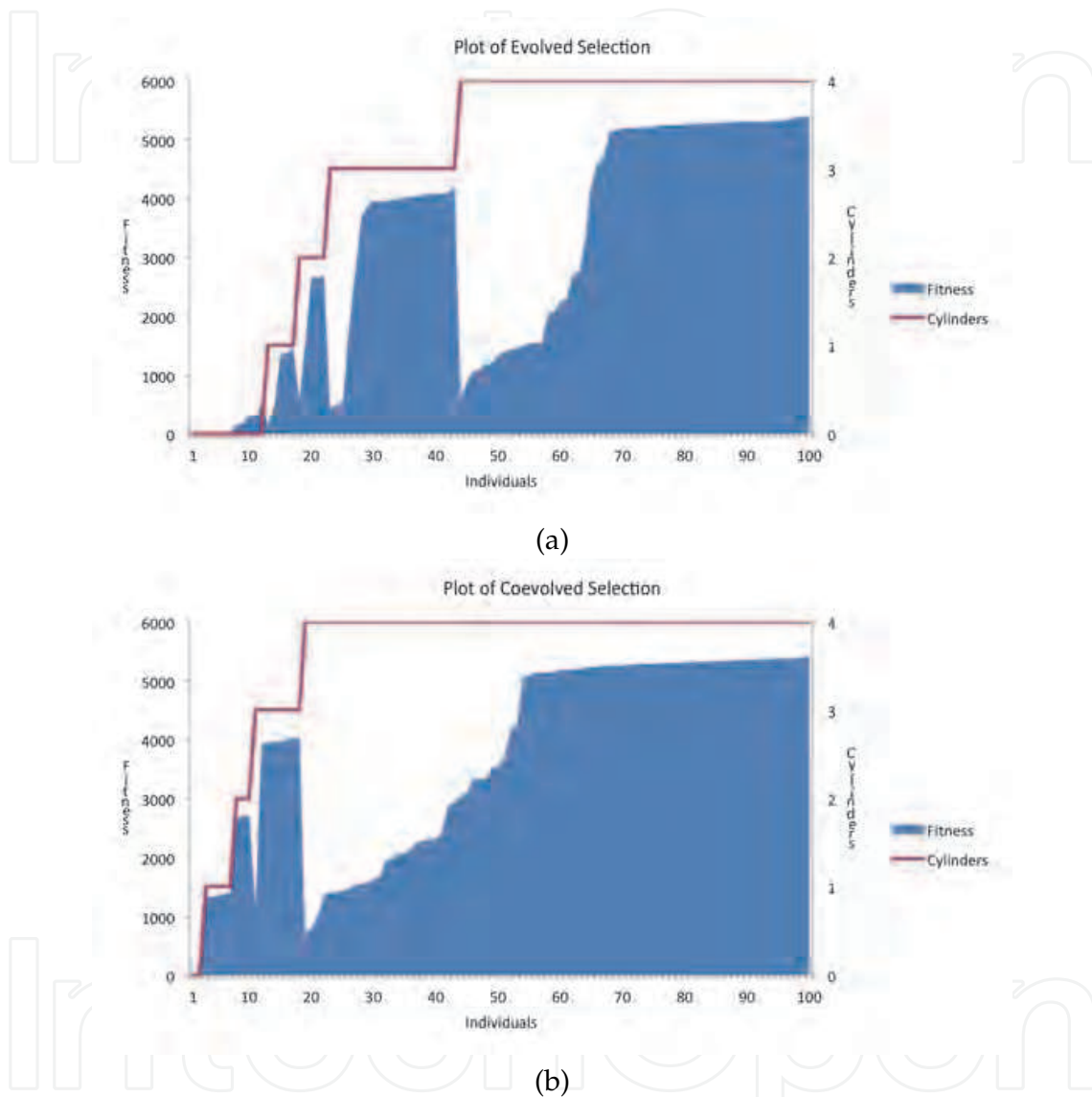


(a)



(b)

Fig. 7. A comparison of one hundred individuals with evolved and coevolved selection is presented. In average an evolved individual collects 3.0 cylinders whereas an evolved individual collects 3.6 cylinders. Moreover, in average evolved selection scores 57% of the highest fitness; in contrast, coevolved-selection scores 69% of the highest evolved fitness. In (a) we notice the first twelve individuals failing to collect cylinders with fifty-seven individuals collecting four cylinders. Whereas in (b) eighty-two individuals accomplish the task of collecting four cylinders and only two individuals not collecting any at all.

## 7. Conclusion

The evolution of central action selection with neural behavior was carried out in this study. Later on, the selection mechanism and neural behavior were coevolved and then compared to evolved selection. The experiments presented in this paper provide an insight of the effects of evolution when optimizing behavior that needs to be coupled within a regular pattern. For example, a disruption of regular selection occurs in an attempt to increase their fitness value as shown in Figures 5 and 6. On the other hand, the use of coevolution constrains candidate solutions to those that maximize the proposed fitness function. Consequently, the maximum fitness is reached when individuals pick up all cylinders in the arena and motivations not affect the localisation of cylinders having to travel less for their collection (Figure 7). Additionally, in order to extend our model we expect to include a component of simulated "dopamine", similar to the one reported in Montes et al. [Montes González et al. (2000)], to regulate behavior through motor commands sent to the Khepera robot. Next, we pretend to analyze motivated behavior at *normal dopamine levels* to see the elicitation of movement (normal selection) and abnormal selection as the result of inducing different levels of simulated dopamine. Furthermore, we expect to develop a prey-predator setup where the prey employs evolvable action selection and the predator a pure evolvable approach both optimized by means of coevolution to study any potential improvements in selection under such a competitive scheme. Also, we may extend our model in a social sense evolving groups of robots. Finally, our approach aims to reduce the number of decisions made by the human designer when evolving both selection and behavior.

## 8. Acknowledgments

## 9. References

Arkin, R. C. (1998). *Behavior-Based-Robotics*, The MIT Press, USA.

Bäck, T. (1996). *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*, Oxford University Press.

Bares, M. & Rektor, I. (2001). Basal ganglia involvement in sensory and cognitive processing. A depth electrode CNV study in human subjects, *Clinical Neurophysiology* 112(11): 2022–2030.

Brooks, R. A. (1986). A robust layered control system for a mobile robot, *IEEE Journal of Robotics and Automation* 2(1): 14–23.

Brooks, R. A. (1989). A robot that walks; emergent behaviors from a carefully evolved network, *Neural Computation* 1(2): 253–262.

Cyberbotics (2011). Webots,Commercial Mobile Robot Simulation Software. URL: *www.cyberbotics.com*

Deb, K. (2001). Multiobjectives optimization using evolutionary algorithms.

Doncieux, S. & Mouret, J. (2010). Behavioral diversity measures for evolutionary robotics, *In IEEE Congress on Evolutionary Computation 2010 (CEC 2010)*.

Ebner, M. (2006). Coevolution and the Red Queen effect shape virtual plants, *Genetic Programming and Evolvable Machines* 7(1): 103–123.

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

Goldberg, D. E. (1989). Genetic Algorithms in search, optimization, and machine learning, *AddisonWesley Reading MA* p. 412.

Holland, J. (1975). *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor.

Kube, C. R. & Zhang, H. (1993). Collective robotics: From social insects to robots, *Adaptive Behavior* 2(2): 189–218.

Kyong Joong, K. & Sung Bae, C. (2001). Robot action selection for higher behaviors with cam-brain modules, *In Proceedings of the 32nd ISR(International Symposium in Robotics)*.

Lapchin, L. & Guillemaud, T. (2005). Asymmetry in host and parasitoid diffuse coevolution: when the red queen has to keep a finger in more than one pie, *Frontiers in Zoology* 2(4): 4.

Maes, P. (1989). The dynamics of action selection, *In IJCAI'89 Proceedings of the 11th international joint conference on Artificial intelligence*, Vol. 2, MI, Detroit, pp. 991–997.

Maes, P. (1990). How to do the right thing, *Connection Science Journal Special Issue on Hybrid Systems* 1(3): 293–325.

Marocco, D. & Nolfi, S. (2007). Emergence of communication in embodied agents evolved for the ability to solve a collective navigation problem, *Connection Science* 19(1): 53–74.

Mcfarland, D. (1993). *Animal Behaviour*, Harlow, Essex, Longman Scientific and Technical.

Mitri, S. & Vogt, P. (2006). Co-evolution of language and behaviour in autonomous robots, *The sixth international conference on the Evolution of Language (Evolang6), World Scientific*, pp. 428–429.

Mondana, F., Franzi, E. & Ienne, P. (1993). Mobile robot miniaturisation: a tool for investigation in control algorithms, *in* S. Verlag (ed.), *Experimental Robotics III, Proceedings of the Third International Symposium on Experimental Robotics*, London, pp. 501–513.

Montes González, F. M. (2007). The coevolution of robot behavior and central action selection, *In IWINAC '07 Proceedings of the 2nd international work-conference on Nature Inspired Problem-Solving Methods in Knowledge Engineering: Interplay Between Natural and Artificial Computation, Part II*, pp. 439–448.

Montes González, F. M., Flandes Eusebio, D. & Pellegrin Zazueta, L. (2008). Action selection and obstacle avoidance using ultrasonic and infrared sensors, *In Frontiers in Evolutionary Robotics*, Book edited by: Hitoshi Iba, I-Tech Education and Publishing, Vienna, Austria, pp. 327–340.

Montes González, F. M. & Marín Hernández, A. (2004). Central action selection using sensor fusion, *In the Proceedings of the fifth Mexican International Conference in Computer Science (ENC'04)*, Baeza-Yates R., Marroquín J.L. and Chávez E (Eds.), IEEE Press, Colima, Mexico, pp. 289–296.

Montes González, F. M., Marín Hernández, A. & Ríos Figueroa, H. (2006). An effective robotic model of action selection, *In R. Marín et al. (Eds.): CAEPIA 2005, LNAI 4177*, pp. 123–132.

Montes González, F. M., Prescott, T. J., Gurney, K., Humphries, M. D. & Redgrave, P. (2000). An embodied model of action selection mechanisms in the vertebrate brain, *In From Animals to Animats 6: Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior*, MA, MIT Press, Cambridge, pp. 157–166.

Montes González, F. M., Santos Reyes, J. & Ríos Figueroa, H. (2006). Integration of evolution with a robot action selection model, *In A. Gelbukh and C. A. Reyes-García (Eds.): MICAI 2006, LNAI 4293*, pp. 1160–1170.

Mouret, J. & Doncieux, S. (2009). Overcoming the bootstrap problem in evolutionary robotics using behavioral diversity, *In Proceedings of the Eleventh Conference on Congress on Evolutionary Computation*, IEEE Press, Piscataway, NJ, USA, pp. 1161–1168.

Nolfi, S. (1997). Evolving non-trivial behaviors on real robots: a garbage collecting robot, *Robotics and Autonomous Systems* 22(3-4): 187–198.

Nolfi, S. & Floreano, D. (2000). *Evolutionary Robotics*, The MIT Press, USA.

Prescott, T. J., Gurney, K., Montes González, F. M., Humphries, M. D. & Redgrave, P. (2002). The Robot Basal Ganglia:Action Selection by and embedded model of the basal ganglia", *In Basal Ganglia VII*, edited by Nicholson L. F. B. and R. L. M. Faull. New York: Kluwer Academic/Plenum Press, pp. 349–358.

Prescott, T. J., Montes González, F. M., Gurney, K., Humphries, M. D. & Redgrave, P. (2006). A robot model of the basal ganglia: behavior and intrinsic processing, *Neural Networks* 19(1): 31–61.

Prescott, T. J., Redgrave, P. & Gurney, K. (1999). Layered control architectures in robots and vertebrates, *Adaptive Behavior* 7(1): 99–127.

Reynolds, C. W. (1994). Evolution of obstacle avoidance behaviour: Using noise to promote robust solutions, *In Advances in Genetic Programming*, MIT Press, Cambridge, MA, USA.

Ridley, M. (2003). *Evolution*, Wiley-Blackwell, USA.

Santos, J. & Duro, R. (2005). *Artificial Evolution and Autonomous Robotics (in Spanish)*, Ra-Ma Editorial, Spain.

Seth, A. K. (1998). Evolving action selection and selective attention without actions, attention, or selection, *In Animals to animats 5: Proceedings of the Fifth International Conference on the Simulation of Adaptive Behavior*, MIT Press, Cambridge, MA, pp. 139–147.

Yamauchi, B. & Beer, R. (1994). Integrating reactive, sequential, and learning behavior using dynamical neural networks, *In From Animals to Animats 3, Proceedings of the 3rd International Conference on Simulation of Adaptive Behavior*, MIT Press, Cambridge, MA, USA, pp. 382–391.
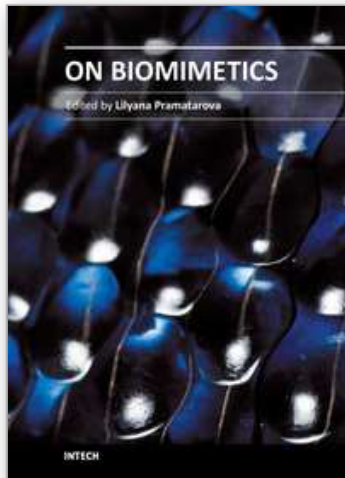
**On Biomimetics**

Edited by Dr. Lilyana Pramatarova

Bio-mimicry is fundamental idea â€˜How to mimic the Natureâ€™ by various methodologies as well as new ideas or suggestions on the creation of novel materials and functions. This book comprises seven sections on various perspectives of bio-mimicry in our life; Section 1 gives an overview of modeling of biomimetic materials; Section 2 presents a processing and design of biomaterials; Section 3 presents various aspects of design and application of biomimetic polymers and composites are discussed; Section 4 presents a general characterization of biomaterials; Section 5 proposes new examples for biomimetic systems; Section 6 summarizes chapters, concerning cells behavior through mimicry; Section 7 presents various applications of biomimetic materials are presented. Aimed at physicists, chemists and biologists interested in biomineralization, biochemistry, kinetics, solution chemistry. This book is also relevant to engineers and doctors interested in research and construction of biomimetic systems.

**How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Fernando Montes-González, Rodrigo Palacios Leyva, Fernando Aldana Franco and Victor Cruz Alvarez (2011). The Coevolution of Behavior and Motivated Action Selection, On Biomimetics, Dr. Lilyana Pramatarova (Ed.), ISBN: 978-953-307-271-5, InTech, Available from: http://www.intechopen.com/books/on-biomimetics/the-coevolution-of-behavior-and-motivated-action-selection

# INTECH
open science | open minds