

A Hybrid Approach in the Development of Behavior Based Robotics

Un Enfoque Híbrido en el Desarrollo de Robótica Basada en el Comportamiento

**Fernando Montes Gonzalez¹, Carlos Alberto Ochoa Ortiz Zezzatti², Luis Felipe Marín-Urías³
and Jöns Sánchez Aguilar⁴**

¹Departamento de Física e Inteligencia Artificial, Universidad Veracruzana
fmontes@uv.mx

²Instituto de Ingeniería y Tecnología (Departamento de Ingeniería Eléctrica y Computación): UACJ
megamax8@hotmail.com

³CNRS-LAAS, Toulouse, France
lfmarin@laas.fr

⁴CIATEC (Centro CONACYT), León de los Aldama; México
jons_sanchez@hotmail.com

Article received on July 31, 2009; accepted on October 29, 2009

Abstract

In this paper we present the development of a method that combines the evolutionary robotics approach with action selection. A collection task is set in an arena where a Khepera robot has to collect cylinders that simulate food. Furthermore, two basic motivations, labeled as ‘fear’ and ‘hunger’, both affect the selection of the behavioral repertoire. In this paper we propose an initial evolutionary stage where behavioral modules are designed as separate selectable modules. Next, we use evolution for optimizing the motivated selection network employed for behavioral switching. Finally, we compare evolved selection with hand-coded selection, which offers some interesting results that support the use of a hybrid approach in the development of behavior-based robotics.

Keywords: Action Selection, Evolutionary Robotics, Behavior-Based Robotics, Bioinspired Algorithms.

Resumen

En este artículo se presenta el desarrollo de un método que combina el enfoque de robótica evolutiva con el de selección de acción. De manera que en una arena se implementa una tarea de recolección para el robot Khepera que debe recoger cilindros simulando comida. Existen dos motivaciones denominadas ‘miedo’ y ‘hambre’ que afectan la selección de módulos conductuales. En este artículo se propone una etapa inicial evolutiva donde se diseñan estos módulos conductuales para que puedan ser elegibles usando selección de acción. Posteriormente se emplea evolución para optimizar la red de selección de acción. Finalmente, se comparan el ajuste de selección obtenido mediante evolución artificial y mediante un diseñador humano, favoreciendo el uso de un enfoque híbrido en el desarrollo de robótica basada en el comportamiento.

Palabras clave: Selección de Acción, Robótica Evolutiva, Robótica Basada en el Comportamiento, Algoritmos Bioinspirados.

1 Introduction

The problem of action selection is identified in ethology as the behavior switching problem. Amongst a behavioral repertoire mostly one behavioral module has to be selected until completion or its execution proves ineffective. In robotics this approach has been widely employed in the behavior-based approach [Arkin, 1998] where an action selection mechanism (ASM) is used to arbitrate between several behavioral modules. Initially, the behavioral modules are developed as separate components that can be integrated by the use of a selector. In this work, these modules can be implemented as neural networks, programming routines, or a mixture of both. As for the tuning of behavior, depending on the chosen implementation this can be developed as hand-coded modules, or 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 these behavioral modules. We aim to achieve this through the interaction of the ASM with the behavioral modules. In our previous work we have employed hand-coded selection to produce regular patterns of behavior [Montes Gonzalez, et al., 2006]; furthermore we have used co-evolution to optimize both

behavior and selection which produce non-regular patterns of behavior [Montes-Gonzalez, 2007]. Additionally, 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. Then, we compare our findings with motivated hand-coded selection.

1.1 Action Selection in the Vertebrate Brain

The action selection problem is related to decision-making whenever a module takes control of the available actuators until is completed or proves ineffective. 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 and Rektor, 2001]. The development of intrinsic basal ganglia circuitry with evolvable behavioral modules has already been implemented in a Khepera robot [Montes, et al., 2007]. Cooperative individuals not only require a society interaction, but the existence of an internal mechanism (e.g. the computational model of the basal ganglia) that is able to mediate amongst 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 et al. (2008) shows how non-standard avoidance can be achieved by extending sensory information through an evolutionary refinement.

1.2 Evolutionary Robotics

Evolutionary robotics employs a quasi-optimal approach to develop autonomous controllers for different kinds of robots. The use of genetic algorithms and neural networks are natural candidates, as the preferred methodology, for developing single evolved neural controllers. These controllers are the result of testing populations of adapted individuals during a refinement process through series of computer-program iterations. Next, pairs or groups of individuals 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 [Lapchin and Guillemaud, 2005]. The latter approach has been identified, as its biological counterpart, as co-evolution 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 co-evolution 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].

2 Evolution and the Design of Behavior

The evolution of the behavior was carried out with a hybrid approach following the evolutionary approach [Nolfi and Floreano, 2000; Santos and Duro, 2005]. Initially modules were evolved in the Webots robot simulator [Cyberbotics, 2010], and later on the modules were further evolved in the real Khepera robot [Mondana et al., 1993]. This robot has been equipped with a ring of eight infrared sensors distributed around the body of the robot and two DC motors 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, 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. In Figure 1 the Khepera robot is shown in a short-walled arena where has to collect blue and red cylinders (color does not make any difference because the camera is not used in these experiments), for this purpose the robot has been equipped with a gripper turret attached to the robot base. Robot 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. The behavioral repertoire is as follows, *cylinder-seek* locates and positions the robot body in front of a cylinder in order to activate *cylinder-pickup* which moves the robot backwards to safely lower the robot arm and then pickup a cylinder; *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; finally *cylinder-deposit* lowers the robot arm, opens the gripper and returns the arm to an upper position.

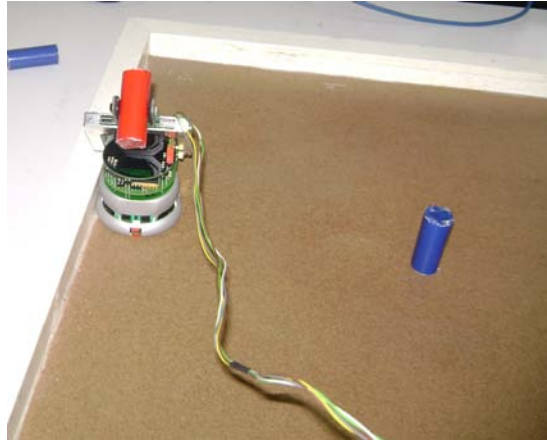


Fig. 1. The Khepera robot set in a short-walled arena

2.1 Exploration Behavior

In our Experiments we use only the infrared sensors of the Khepera robot to transverse the arena. Thus, we chose to implement exploration behavior as neural controllers. Behavioral modules implemented in such way correspond to *wall-seek*, *corner-seek* and *cylinder-seek*. They 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, four 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 ± 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. 1-3).

The weights of each neural network are directly encoded into a vector \mathbf{w} of 32 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. Then, 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 25 seconds using the fast-speed mode of the Webots simulator. 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

$$f_{c1} = \sum_{i=0}^{3500} \text{abs}(ls_i) * (1 - \sqrt{ds_i}) * (1 - \max_ir_i) \quad (1)$$

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

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

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

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

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*).

2.2 Gripper-Handling 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.

3 Motivated Action Selection Mechanism

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. Brooks, 1986; Maes 1989; Arkin, 1998]. 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 Gonzalez 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-Gonzalez, et al., 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-Gonzalez and 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.

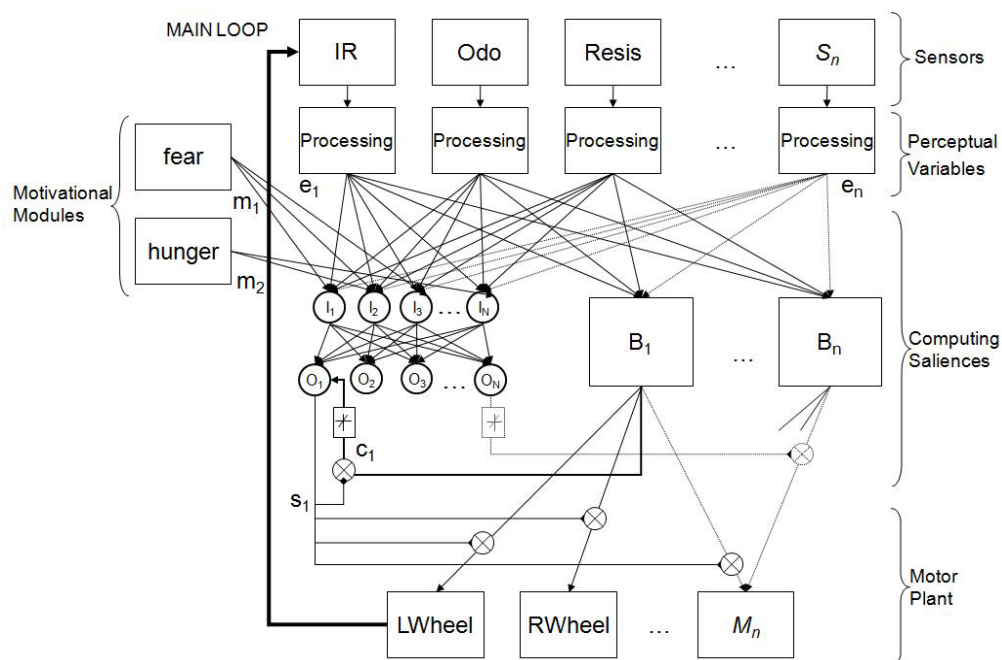


Fig. 2. CASSF produces Action Selection by converting sensory input into perceptual variables; next salience is calculated and then motor output is sent to the motors

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. In addition, CASSF is an effective action selection mechanism [Montes-Gonzalez, et al., 2006] that is centralized and presents sufficient persistence to complete a task. The implementation of tasks such as foraging can be carried out by determining a set of behavioral patterns that can be integrated in time to complete such a task. Furthermore, selection parameters of this model have been optimized by the use of evolution. The adjustment of selection parameters and behavior has been optimized by co-evolution in CASSF as described in [Montes-Gonzalez, 2007]. In this 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_i) from behavior B_i to the output neuron O_i 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 (Figure 2).

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-take-all* manner.

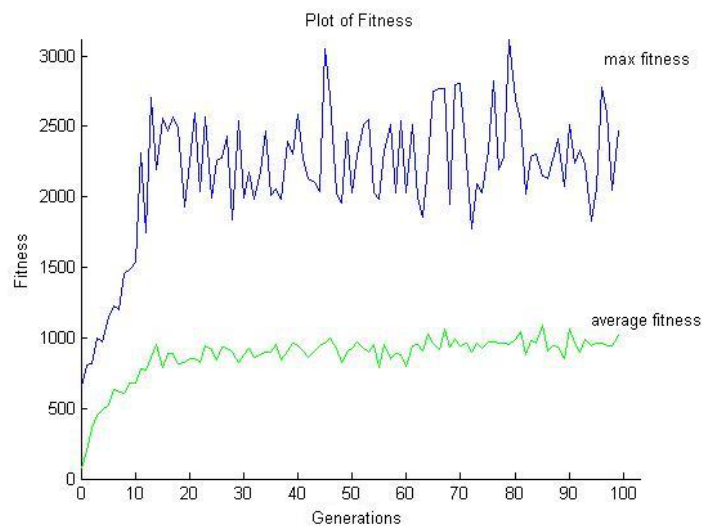


Fig. 3. Population fitness (Eq. 4) is plotted across one hundred generations

In this paper artificial evolution 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 the decision network weights. However, these weights depend on the input of the context vector (\mathbf{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 ($\mathbf{e} = [e_w, e_g, e_c, e_r, e_w, e_g, e_c, e_r \in \{1, -1\}]$). Then, five different behavioral modules return a current busy-status (\mathbf{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-seeking, cylinder-pickup, wall-seeking, corner-seeking, and cylinder-deposit respectively. The motivational vector is composed by $\mathbf{m} = [m_f, m_h], -1 < \mathbf{m}_i < 1$. The salience (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 $K_w = 0.75$.

Evolution was carried out as previously described in section 2.1. The fitness formula for the evolution of the weights of the decision network was

$$f_{c4} = (K_1 * cwfactor) + (K_2 * f_{c2}) + (K_3 * pkfactor) + (K_4 * dpfactor) \quad (4)$$

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 locate corners and walls in the arena ($cwfactor$). On the other hand, the fitness formula also rewards locating cylinders (f_{c2}), their collection inside the arena ($pkfactor$), and their release near the outside walls ($dpfactor$). 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 2500. In Figure 3 we present the average fitness, and its maximum individual fitness, for over 100 generations.

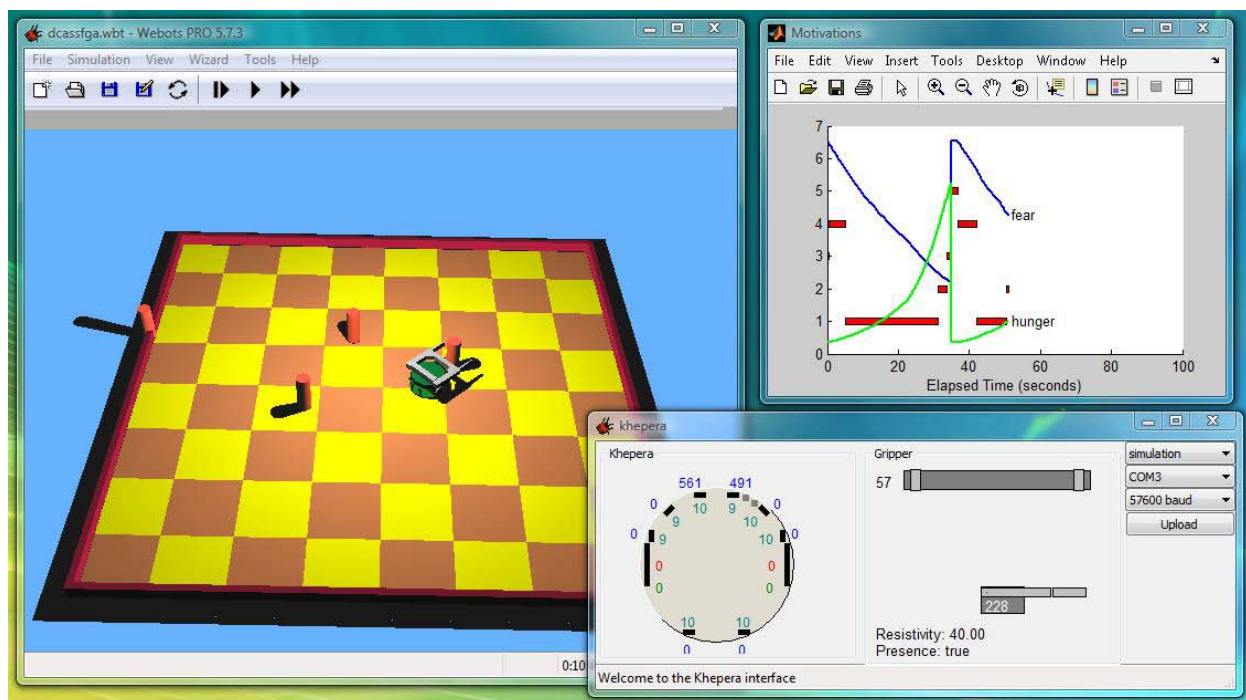


Fig. 4. Action selection is modulated by internal motivational variables like fear and hunger

4 Experiments and Results

The foraging task was set in an arena with four cylinders as simulated food. Figure 4 presents the robot simulator window and the window for the gripper and infrared sensors. In the motivations window, the blue line corresponds to 'fear' and the green line represents 'hunger'. Next, final behavior is transferred to the real robot for further optimization. The use of selection with hand-coded parameters and evolved behavioral modules is shown in Figure 5 where we notice that although four cylinders were collected only three were delivered right next to the corner. An individual such as this presents a regular grasping-depositing pattern (behavioral pattern selection of 1-2-3-4-5). Selection behavior is also summarized as elementary statistics. The labels in Table 1 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.

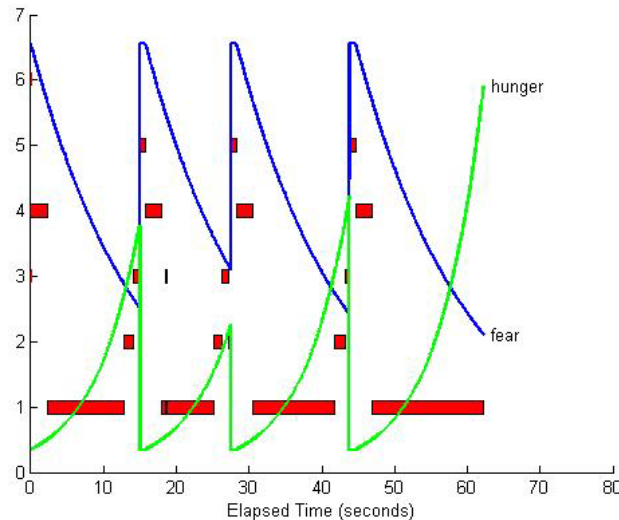


Fig. 5. Ethogram for a run of typical hand-coded behavior selection; modules are numbered as 1-cylinder-seek, 2-cylinder-pickup, 3-wall-seek, 4-corner-seek, 5-cylinder-deposit and 6-no action selected

Table 1. Hand-coded selection (Figure 5) is summarized as elementary statistics

Behavioral Modules	Freq	Latency	TotDur	TotDur%	Mean	StdDev	StdErr	MinDur	MaxDur
none	1.00	0.00	0.02	0.03	0.02	0.00	0.00	0.02	0.02
cylinder-seek	5.00	2.23	44.02	70.69	8.80	5.66	2.53	0.45	15.37
cylinder-pickup	4.00	12.81	4.09	6.56	1.02	0.53	0.26	0.25	1.39
wall-seek	5.00	0.02	2.86	4.58	0.57	0.42	0.19	0.03	1.05
corner-seek	5.00	0.05	8.91	14.30	1.78	0.99	0.44	0.02	2.26
cylinder-deposit	3.00	14.98	2.39	3.83	0.80	0.02	0.01	0.78	0.81
Total	23.00	0.00	62.28	100.00	2.71	4.13	0.86	0.02	15.37

On the other hand, staged evolution of behavior and selection is shown in Figure 6 where a standard grasping-depositing pattern is not easily observed. Because of opportunism exhibited 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 an additional fifth cylinder is picked up and released making a total of five collected cylinders (the rest are fallen attempts). As a result, a semi-regular behavioral pattern selection of 1-2-3-5 can be observed in this graph. We notice from elementary statistics in Table 2 that although cylinder-pickup was selected 11 times, cylinder-deposit was only selected 5 times. The latter is due to the fact that fallen attempts to collect cylinders are accounted as triggered behavior. Though, an already released cylinder is immediately grasped and then released again around second 30 (in Figure 6). Additionally, evolution avoids the use of corner-seek and that is reason for its latency to be shown as the total elapsed time even though selection never occurred. As the result of evolution the fitness function is shaping selection by the optimization of behavior in time and in the physical environment.

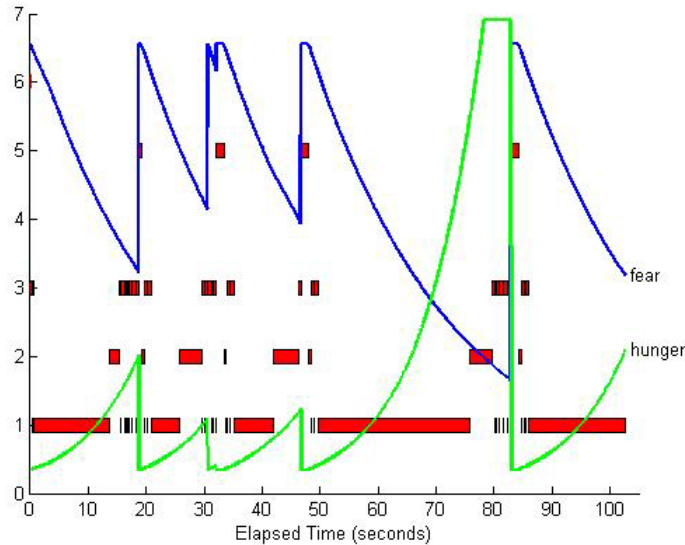


Fig. 6. A typical run for evolved selection; the behavioral modules are numbered as 1-cylinder-peek, 2-cylinder-pickup, 3-wall-peek, 4-corner-peek, 5-cylinder-deposit and 6-no action selected. A standard grasping-depositing pattern is not easily observed on this ethogram

Table 2. Elementary statistics for evolved selection

Behavioral Modules	Freq	Latency	TotDur	TotDur%	Mean	StdDev	StdErr	MinDur	MaxDur
none	1.00	0.00	0.03	0.03	0.03	0.00	0.00	0.03	0.03
cylinder-peek	49.00	0.45	68.94	67.09	1.41	4.79	0.68	0.02	26.07
cylinder-pickup	11.00	13.62	15.60	15.18	1.42	1.74	0.53	0.02	4.27
wall-peek	44.00	0.03	13.85	13.48	0.31	0.25	0.04	0.02	0.76
corner-peek	0.00	102.76	0.00	0.00	NaN	NaN	NaN	0.00	0.00
cylinder-deposit	5.00	18.74	4.34	4.22	0.87	0.60	0.27	0.02	1.33
Total	110.00	0.00	102.76	100.00	0.93	3.27	0.31	0.00	26.07

In our work 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-peek, cylinder-pickup, wall-peek, corner-peek, 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 the disruption of 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 (after second 50 in Figure 6). Finally, the fitness of the agent solving the foraging task is an additional factor that alters the order in the selection of behavior. In general a regular behavioral pattern is easily observed for hand-coded selection whereas for evolved selection the use of corner-peek is avoided in order to optimize time selection.

In Figure 7 we present a comparison of hand-coded-selection/evolved-behavior and staged evolved-selection/behavior. Here we notice that in average a hand-coded individual collects one cylinder whereas an evolved individual collects three cylinders. Furthermore, in average hand-coded selection scores 32% of the highest hand-coded fitness; in contrast, evolved-selection scores 63% of the highest evolved fitness. In (a) we notice the first forty individuals failing to collect cylinders with six individuals collecting four cylinders. Whereas in (b) fifty-nine individuals accomplish the task of collecting four cylinders and only eleven individuals not collecting any at all. 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. Consequently, in the same figure we observe an improvement on the fitness and the collection of cylinders of evolved selection in comparison to hand-coded selection with evolved behavior.

5 Discussion and Future Work

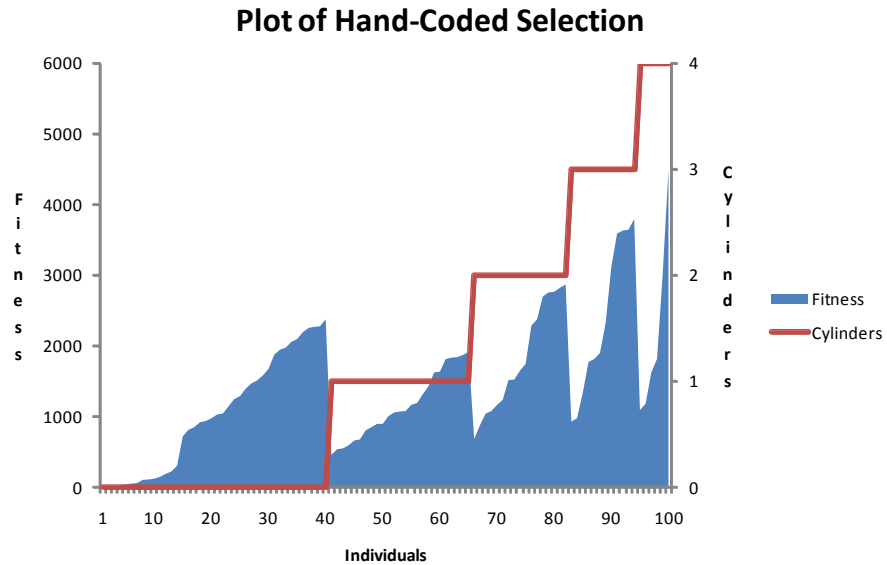
For our discussion it is important to remember that there is evidence of central selection in the vertebrate brain [Prescott et al., 1999], particularly at the basal ganglia buried under the cortex. It is important to notice that these structures receive information from several different regions of 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 developed a hybrid 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 simulated fear and hunger to affect the results of selection. Therefore, selection arbitrates amongst 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. In order to extend our model we expect to include a component of simulated 'dopamine', similar to the one reported in Montes 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 pretend 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 co-evolution to study any potential improvements in selection under such a competitive scheme.

6 Conclusion

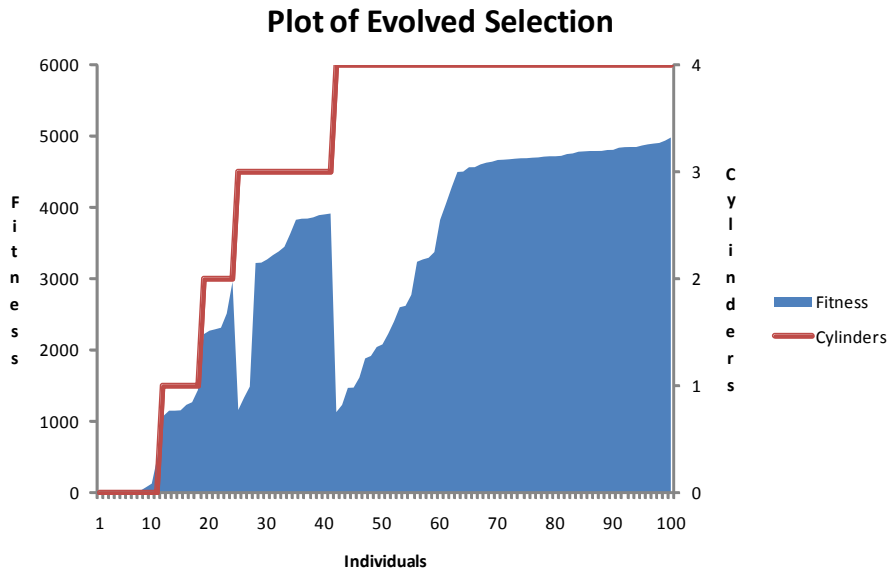
The evolution of central action selection with neural behavior was carried out in this study. Later on, both the selection mechanism and neural behavior were further evolved in two separate stages and then compared to hand-coded selection with evolved neural behavior. The experiments presented in this 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 Figure 6. On the other hand, the use of evolution constrains candidate solutions to those that maximize the proposed fitness function. Consequently, the maximum fitness is reached when evolved instead of hand-coded selection is employed (Fig. 7). Finally, our hybrid approach aims to reduce the number of decisions made by the human designer when evolving both selection and behavior.

Acknowledgments

The authors would like to thank Nicandro Cruz Ramírez for his insightful comments on the preparation of this paper.



(a)



(b)

Fig. 7. A comparison of one hundred individuals after the evolution of hand-coded and optimized selection

References

1. **Arkin, R. C. (1998).** *Behavior-Based Robotics*. Cambridge, Mass: The MIT Press.
2. **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.

3. **Brooks, R. A. (1986).** A Robust Layered Control System for a Mobile Robot. *IEEE Journal of Robotics and Automation*, 2(1), 14-23.
4. **Brooks, R. A. (1989).** A robot that walks; emergent behaviors from a carefully evolved network. *Neural Computation*, 1(2), 253-262.
5. **Cyberbotics (2010, February 2).** Webots, Commercial Mobile Robot Simulation Software. Retrieved from <http://www.cyberbotics.com>
6. **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). Retrieved from <http://www.frontiersinzoology.com/content/2/1/4>
7. **Maes, P. (1989).** The Dynamics of Action Selection. *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence, IJCAI-89*, Detroit, MI, 991-997.
8. **Mcfarland, D. (1994).** *Animal Behaviour*. Harlow, Essex: Longman Scientific and Technical.
9. **Mondada, F., Franzi, E. & Jenne, P. (1993).** Mobile robot miniaturisation: a tool for investigation in control algorithms. *Third International Symposium on Experimental Robotics. Lectures Notes in Control and Information Science*, 200, 501-513
10. **Montes Gonzalez, F., Prescott, T. J., Gurney, K., Humphries, M. & Redgrave, P. (2000).** An embodied model of action selection mechanisms in the vertebrate brain. In J. A. Meyer (Ed.), *From Animals to Animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behavior*, Paris, France, 157-166.
11. **Montes-González, F. M., & Marín-Hernández, A. (2004).** Central Action Selection using Sensor Fusion. In R. Baeza-Yates, J. L. Marroquín & E. Chávez (Eds.), *Proceedings of the Fifth Mexican International Conference in Computer Science, ENC'04*, Colima, México, 289-296.
12. **Montes-González, F. M., Marín Hernández, A. & Ríos Figueroa, H. (2006).** An Effective Robotic Model of Action Selection. *11th Conference of the Spanish Association for Artificial Intelligence, CAEPIA 2005. Lecture Notes in Artificial Intelligence*, 4177, 123-132.
13. **Montes-González, F., Santos Reyes, J., & Ríos Figueroa, H. (2006).** Integration of Evolution with a Robot Action Selection Model. *5th Mexican International Conference on Artificial Intelligence, MICAI 2006. Lecture Notes in Artificial Intelligence*, 4293, 1160-1170.
14. **Montes-Gonzalez, F., Prescott, T. J. & Negrete-Martinez, J. (2007).** Minimizing human intervention in the development of basal ganglia-inspired robot control. *Applied Bionics and Biomechanics*, 4 (3), 101-109.
15. **Montes-Gonzalez, F. (2007).** The Coevolution of Robot Behavior and Central Action Selection. *Second International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2007. Lecture Notes in Computer Science*, 4528, 439-448.
16. **Montes-Gonzalez, F., Flandes-Eusebio, D., & Pellegrin-Zazueta, L. (2008).** Action Selection and Obstacle Avoidance using Ultrasonic and Infrared Sensors. In Hitoshi Iba (Ed.), *Frontiers in Evolutionary Robotics (327-340)*. Vienna, Austria: I-Tech Education and Publishing.
17. **Nolfi, S. & Floreano, D. (2000).** *Evolutionary Robotics*. Cambridge, Mass: The MIT Press.
18. **Prescott, T. J., Redgrave, P. & Gurney, K. (1999).** Layered control architectures in robots and vertebrates. *Adaptive Behaviour*, 7(1), 99-127.
19. **Prescott, T. J., Gurney, K., Montes-Gonzalez, F., Humphries, M. & Redgrave P. (2002).** The Robot Basal Ganglia: Action selection by and embedded model of the basal ganglia. In L. F. B. Nicholson & R. L. M. Faull (Eds.), *Basal Ganglia VII (349-358)*. New York: Kluwer Academic/Plenum Press.
20. **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.
21. **Ridley, M. (2003).** *Evolution*. (3rd ed.). Carlton, Victoria: Wiley-Blackwell.
22. **Santos, J. & Duro, R. J. (2005).** *Evolución Artificial y Robótica Autónoma*. Madrid: Ra-Ma Editorial.



Fernando Montes Gonzalez studied the Bachelor degree in computer science at the Universidad Veracruzana in Xalapa, Veracruz, Mexico. He obtained a M. Sc. Degree in Computer Science: Artificial Intelligence from the Universidad Veracruzana and a Ph.D. degree from the University of Sheffield in England, UK. His research interests include action selection, evolutionary robotics, behavior-based robotics, and hybrid artificial intelligence systems.



Carlos Alberto Ochoa Ortiz Zezzatti (Bs'94 - Eng. Master'00 – Ph.D.'04 - Postdoctoral Researcher'06 & Industrial Postdoctoral Research'08). He has written one book and seven chapters in books related to AI. He has supervised seven Ph.D. theses, eleven Master theses and twenty seven Bachelor theses. He participated in the organization of conferences such as HAIS'07, HAIS'08, ENC'06, ENC'07, ENC'08 and MICAI'09. His research interests include evolutionary computation, natural processing language and social data mining.



Luis Felipe Marin Urias has a Bachelor degree in computer science from the Universidad Veracruzana. He studied a Master degree in computer science: Artificial Intelligence at the same university. He obtained his Ph.D. in the field of human-robot interaction at the LAAS-CNRS and the University of Toulouse III in France. His research interests include evolutionary robotics and hybrid artificial intelligence systems.



Jöns Sanchez Aguilar (Bs'00 - Eng. Master'02 both from the Instituto Tecnológico de Querétaro and the Ph.D.'07 from the CIATEC-Conacyt Research Center). His research field is the mathematical modeling and optimization of process and products using response surface methodology and genetic algorithms. Furthermore, he is a professor of the postgraduate program at the CIATEC.