

Selective Attention Enables Action Selection: Evidence From Evolutionary Robotics Experiments

Giancarlo Petrosino, Domenico Parisi, Stefano Nolfi

Institute of Cognitive Sciences and Technologies

National Research Council (CNR-ISTC)

Via S. Martino della Battaglia, 44, Rome, Italy

giancarlo.petrosino@istc.cnr.it, domenico.paris@istc.cnr.it, stefano.nolfi@istc.cnr.it

Abstract

In this paper we investigate whether selective attention enables the development of action selection (i.e. the ability to select among conflicting actions afforded by the current agent/environmental context). By carrying out a series of experiments in which neuro-robots have been evolved for the ability to forage so to maximize the energy that can be extracted from ingested substances we observed that effective action and action selection capacities can be developed even in the absence of internal mechanisms specialized for action selection. However, the comparison of the results obtained in different experimental conditions in which the robots were or were not provided with internal modulatory connections demonstrate how selective attention enables the development of a more effective action selection capacity and of more effective and integrated action capacities.

1. Introduction

To survive and reproduce animals must engage in many different behaviours. In some cases, the selection of the behaviour to be currently displayed naturally follows from the state of the external sensors encoding information about the external environment and the internal sensors encoding information about the organism's needs. But in many circumstances the internal and external environment present animals with many opportunities and demands for actions. For example, the perception of water offers an opportunity to satiate thirst but the concurrent perception of a predator demands an appropriate defensive response. Alternative actions relying on the same motor effectors (e.g. approaching water and flying away from the predator) cannot be performed at the same time. Thus, one fundamental capacity that organisms should display consists in action selection (Seth, Prescott and Bryson, 2012), i.e., the capacity to select between alternative actions afforded by the current organism/environmental context. In that respect, action selection should be considered as an instance of the more general capacity of producing multiple integrated behaviours. Action selection can also be interpreted as motivational choice, i.e., choosing which motivation among one's many different motivations should currently control one's behaviour (Parisi and Petrosino, 2010; Ruini et al 2012).

Although action selection and more generally action integration has been and is actively studied in different disciplines (ranging from ethology, neurobiology and psychology to artificial intelligence and robotics) the question of whether it can be simply considered as the emergent result of the interaction of multiple sensory-motor processes or whether it requires specialized neural mechanisms is still controversial (Seth, 1998; Prescott, 2007; Seth, Prescott and Bryson, 2012).

According to many authors (Neumann, 1987; Tipper, Howard, & Jackson, 1998; Castiello, 1999; Cisek, 2012) action selection is enabled by selective attention: i.e. the process by which organisms select (and process) a subset of available information by directing their sensory organs toward specific stimuli (overt attention) and/or by internally focusing on one part of the available sensory stimuli (covert attention). This hypothesis is in line with ecological accounts (Gibson, 1979) that tend to emphasize the fact that perceptions can flow directly into actions with little or no intention to act (Gibson, 1979; Duncan-Johnson and Koppell, 1981; Goldberg and Segraves, 1987; Miller and Hackley, 1992). For different accounts that rather tend to emphasize the mediation role of internal processes, such as the perceptual control theory (Powers, 1973), selective attention does not necessarily represent a prerequisite for action selection.

In humans, the relation between selective attention and action selection is supported by experiments in which human subjects are asked to reach and grasp a target object in the presence of distractor objects. The analysis of these experiments indicate that when the attention is allocated to an object, only the perceptual properties of the attended object that are relevant for the execution of the action (i.e., its volumetric properties and its spatial properties with respect to the hand's frame of reference) are computed (Howard and Tipper, 1997; Tipper et al, 1997; Jackson et al. 1995; Chieffi et al. 1993; Castiello, 1996; Bonfiglioli and Castiello, 1998, Tipper et al, 1998). On the contrary, when attention is also driven toward a distractor object, the characteristics of the distractor object interfere with those of the target object during the execution of the action (Castiello, 1999). Moreover, Riddoch and collaborators has shown that the presence of distractor objects strongly interfere in reaching-to-grasp actions directed toward a target object in a brain-injured patient displaying attentional deficits (Riddoch et al., 2000).

As far as we know, the relation between selective attention and action selection in invertebrate have not been investigate systematically. However the modulation of sensory processes occurring during the execution of specific actions have been observed in crayfish (El Manira et al, 1991), stick insect (Bassler and Buschges, 1998), cricket (Gebhart and Honnegger, 2001), locust (Buschges and Wolf, 1999) and in many other species.

In this paper we investigate the relation between selective attention and action selection through a synthetic method, i.e. by evolving neuro-robots that need to engage in several activities to survive and reproduce. More specifically we investigate whether the availability of a top-down inhibitory mechanism that enable the active suppression of non-relevant stimuli support the development of more effective action selection capabilities.

The results support the hypothesis that selective attention facilitates/enables action selection. In the next section we discuss the method and its relation with respect to the state of the art. In section 3 we present the experimental scenario. In section 4 we present the obtained results. Finally in section 5 we draw our conclusions.

2. Method

One important consideration in modelling action selection is that in natural systems behaviours are not atomic elements but rather processes that emerge from the agent/environment interactions. They cannot be interpreted as properties or elements internal to the agent or manifestation of properties/mechanisms that are internal to the agent and, consequently, action selection cannot be considered as a process that occurs only inside the agent. In contrast with this, the majority of models of action selection that have been elaborated so far (including embodied models) assume that the units of selection consist of internal correlates of behaviours: a set of behavioural layers or

modules located inside the agent, possibly organized in a hierarchical manner, which are able to produce specific types of behaviours (Brooks, 1986; Maes, 1990; Tyrrel, 1993; Bryson, 2000; Montes Gonzales et al, 2000).

In line with the approach followed by Anil Seth (1988; 2012), the model presented in this paper avoids this implicit assumption by using embodied and situated agents that develop their skills by adapting to their task/environment. The evolution process is driven by an evaluation criterion that does not specify the specific behaviours and behavioural skills that should be realized and selected and/or the way in which they should be produced and, therefore, leaves the robots free to exploit the properties emerging from the robot/environment interactions.

A second important consideration concerns the dynamical and contextual nature of behaviours. The embodied and situated nature of behaviour implies that the actions should be flexible enough to achieve a specific outcome in varying conditions (e.g. a reaching behaviour should allow a robot to bring its hand on objects located in varying relative positions) and to be effectively combined with other actions (e.g. reaching and grasping behaviours should be realized in a way that enables their combination over time to pick up a distant object). This flexibility clearly impacts on action selection. More precisely, the manner in which actions are realized at a certain developmental stage constitutes an adaptive constraint on how actions should be selected. Conversely, the way in which actions are selected at a certain development stage constitutes an adaptive constraint on the manner in which actions should be realized.

The use of an adaptation process in which the agents are left free to determine how action are realized and selected and in which these two capabilities are co-adapted allows us to take into account the effect of the interdependencies between action production and action selection. Indeed, as we will see, our results demonstrate a clear interdependence between the way in which actions are selected and realized.

Another implication of the dynamical and emergent nature of behaviour is constituted by the lack of clear cues that can be used to partition the overall behaviour produced by an agent into a set of more elementary behaviours. Behaviour identification and partitioning comes as a natural process for a human observing a behaving agent. On the other hand, the use of human observers for the identification of behavioural units would clearly introduce biases caused by the external perspective and by individual variability. This issue touches a deep philosophical question on whether behaviours can be considered real entities eligible for scientific analysis or subjective entities that only exist in the eyes of the observed (for a synthetic discussion, see Prescott, 2008). This paper does not provide an answer the subjective-objective question. On the other hand the use of a synthetic method in which all variable and parameters are known and in which the selection criterion that drive the evolutionary process is formalized significantly reduces the problems caused by this intrinsic difficulty. Indeed it allowed us to operationalize the process through which the overall behaviour exhibited by the robots is divided into behavioural units.

Finally, we will investigate whether the availability of top-down connections and of neuro-modulatory mechanism, that can allow the robots to reduce or eliminate the processing of non-relevant stimuli, can facilitate the acquisition of the required behavioural capacity and the acquisition of the capacity to appropriately select among them. For a related connectionist model of negative priming and of the neurophysiological mechanisms behind attention, see Schrobsdorff, Ihrke, Behrendt, Hasselhom and Hermann (2012).

3. The experimental setup

To investigate the issues described above we have evolved neuro-robots for the capacity to survive and reproduce. The robots need to ingest various substances to remain alive but they have also to take into account what their body needs at any given time. The need to collect different food sources and the need to take into account the current state of their body which dynamically changes over time require the development of several behavioural skills as well as the ability to select the

right behaviour at the right time. The chosen scenario does not aim to model a specific animal species or environmental niche. However, it is designed so to take into account the fact that natural organisms: (i) are situated in rich environments containing different type of objects/resources; (ii) are required to display several behaviours (including mutually exclusive and/or interfering behaviour); (iii) are required to face different adaptive constraints (e.g., forage, save energy, avoid risks); (iv) are characterized by internal processes (e.g., metabolic processes) and states that co-determine their behaviour; (v) are provided with sensors that provide egocentric, limited, and noisy information; (vi) are selected on the basis of how their overall behaviour corresponds to the adaptive constraints and not on the basis of whether they display pre-determined behaviours and/or on the basis of how they select among them.

3.1 The robot and the environment

The agents used in our experiment are simulated e-Puck robot (Mondada et al., 2009; <http://www.e-puck.org/>), i.e., small circular wheeled robots provided with 8 infrared sensors uniformly distributed around the robot body, a VGA camera located on the frontal side of the robot, and two motors controlling the speed of the robot's two wheels. The robot is situated in a square arena of 150x150cm surrounded by walls and containing different types of food (red, green, and blue food elements, see Figure 1). More precisely, the environment contains: (a) five green cylindrical objects (diameter 25mm) made of green edible substance, (b) one blue cylindrical object (diameter 25mm) that disperses a blue edible substance over a nearby circular area, (c) one red cylindrical object (diameter 50mm) that disperses a read edible substance over the nearby area. Green objects contain one unit of green substance which is automatically ingested by the robot as soon as the distance between the robot and the object is below 3mm (green objects disappear from the environment once they are ingested). In addition, robots ingest 0.05 and 0.1 units of blue or red substance, respectively, each time step spent near the blue and red objects (i.e. within a distance of 150mm and 60mm, respectively). The quantity of the green, blue and red substances within the robot body, that is initially 0.0, increases up to its maximum level of 1.0 as a consequence of substance ingestion, and decrease of 0.00125, 0.00125, and 0.0 each time step as a consequence of substance degradation over time for green, blue and red substances, respectively (i.e. the red substance does not degrade).

The robot's "stores" ingested green, blue and red substances inside its body and is provided with idealised internal organs that synthesize energy from the ingested food substances through a "metabolic" process. More precisely, for each time step in which the quantity of both green and blue substances overcomes a threshold of 0.5, the robot's energy level is increased by the following quantity:

$$E_t = 0.0025 * (1 + QR) \quad (1)$$

where E is the synthesized energy at time step t (i.e. during 100ms), 0.0025 is a constant, and QR is the current quantity of red substance. This implies that to produce energy the robot should periodically ingest green and blue substances and, since the availability of red substance maximize the energy that can be extracted from blue and green substances, the robot should ingest a 1.0 quantity of red substance as earlier as possible. The robot's body and the motors of the robot consume energy according to the following equation:

$$E_t = - (0.0005 + 0.0015 * S) \quad (2)$$

where 0.0005 is the constant amount of energy consumed at each time step t (100 ms) by the robot's body to remain operational, 0.0015 is the maximum amount of energy consumed by the robot's motors, and S is the average absolute speed of the two wheels normalized in the interval $[0, 1]$. This

means that the robot consumes a small fixed amount of energy and a variable amount of energy proportional to the speed of its motors in any time step.

At the beginning of each trial the energy level is initialized to 0.0. After 1000 time steps in which the energy level remains at a null value, the robot is not allowed to move anymore for the rest of the trial.

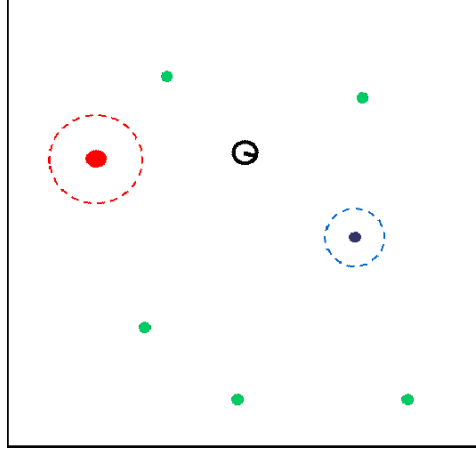


Figure 1. The robot and the environment. The black circle represents the robot. The black square represents the walls surrounding the robot's environment. The red, green, and blue disks represent the red, green, and blue food elements, respectively. The red and blue dashed circles indicate the areas within which the robot absorbs the corresponding substances.

3.2 The robot's nervous system

The nervous system of the robot is constituted by a recurrent neural network with 17 sensory neurons, 4 proprioceptors, 6 internal neurons, and 3 motor neurons (Figure 2).

The sensory layer is constituted by: (i) eight infrared sensory neurons that encode the state of the corresponding robot's infrared sensors, (ii) three groups of three visual sensory neurons that encode the percentage of red, green, and blue light detected by the camera on the left, frontal, and right sector of the visual field, respectively (the camera has a field of view of 36° that is divided into three sectors of 12° each), and (iv) four proprioceptors neurons that encode the current quantity of red, green, and blue substances contained in the robot body's and the current level of robot's energy. The last four neurons thus encode information related to the robot's needs that is particularly relevant for action selection. The state of all sensory neurons is normalized in the range $[0.0, 1.0]$. A white noise in the range $[-0.05, 0.05]$ is added to the activation level of the infrared and visual sensory neurons to take into account the variability of sensory measures in the real environment. The experiment has been carried out by using Evorobot* (Nolfi & Gigliotta, 2010), a software tool that has been successfully used to port simulation results in hardware.

The motor layer is constituted by two motor neurons that encode the speed of the left and right wheels normalized in the range $[-maxs, maxs]$ and by one motor neuron that determines the maximum speed of the two wheels ($maxs$) in the range $[0, 2\pi]$ rad/sec. In other words, the third neuron regulates the speed of the wheels independently from the state of the first two motor neurons. This design choice has been done to enable the robots to regulate their speed more easily.

The activity of the internal and motor neurons is computed on the basis of the following equation:

$$A_j = \vartheta_j + \sum_{i,j=1}^N w_{ij} O_i \quad (3)$$

with A_j being the activity of the j th neuron, ϑ_j the bias of the j th neuron, w_{ij} the weight from the i th to the j th neuron, O_i the output of the i th neuron.

The output of the sensory, internal, and motor neurons is set on the basis of the following three corresponding equations:

$$O_j = s_j g_j \quad (4)$$

$$O_{j(t)} = (\tau_j O_{j(t-1)} + (1 - \tau_j)(1 + e^{-A_{j(t)}})^{-1}) g_{j(t)} \quad (5)$$

$$O_j = (1 + e^{-A_j})^{-1} g_j \quad (6)$$

with s_j being the state of the corresponding sensors normalized in the range [0.0, 1.0], τ_j being the time constant of the j th neuron (which is a parameter in the range [0.0, 1.0] in the case of internal neurons and a constant value set to 0.0 in the case of other neurons), $O_{j(t-1)}$ being the output of the neuron at the previous time step, g_j being the gain of the j th neuron (which is set to the output of the corresponding modulatory neuron at time $t-1$ in the case of modulated neurons and to 1.0 in the case of the other neurons), and A_j being the activity of the j th neuron.

The experiments have been replicated in different experimental conditions that vary with respect to the way in which neurons are connected and with respect to the presence of regulatory circuits (see section 4). As we will see, the architecture shown in Figure 2c facilitates the emergence of more effective solutions.

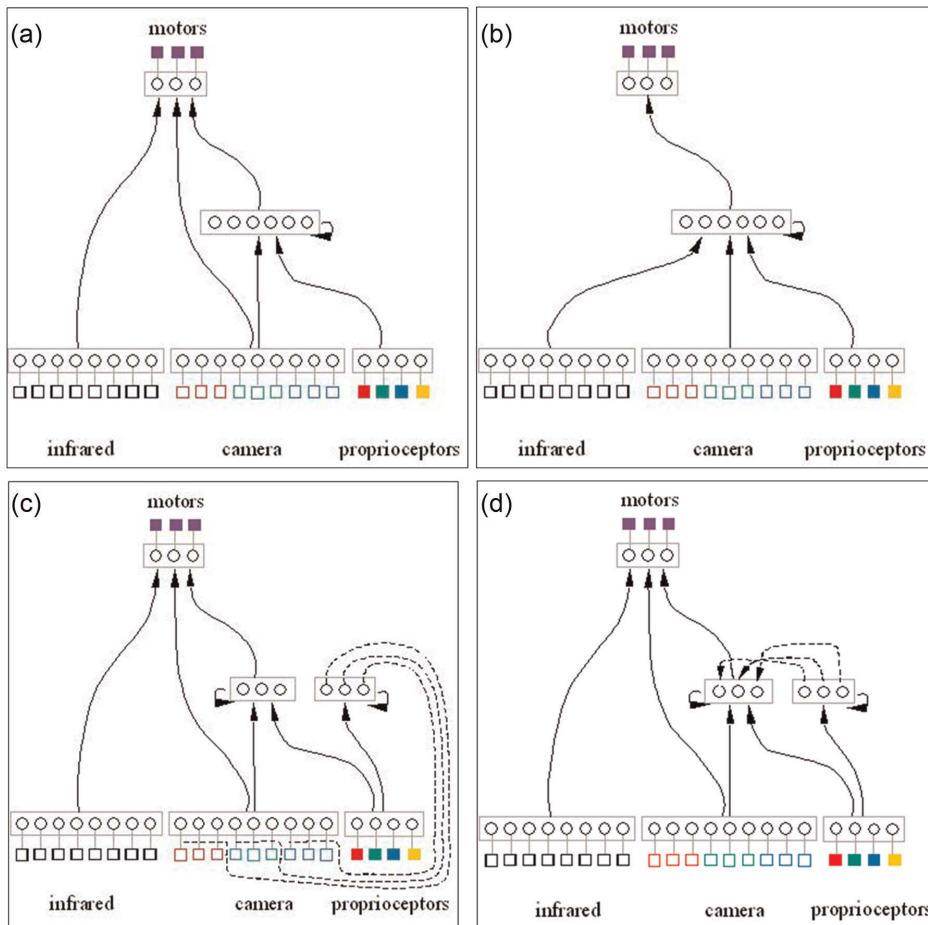


Figure 2. The robot's nervous system relative to the four different experimental conditions (a) A, (b) B, (c) C and (d) D (see section 4.1). (a) and (b): standard architecture with recurrent internal neurons. (c) and (d): architecture including a modulatory circuit constituted by three internal neurons that determine the gain of the 9 sensory neurons encoding visual information (c) or the gain of the three hidden neurons (d). Squares represent sensors and motors. Circles represent neurons. Rectangle represent blocks of neurons. Solid arrows represent connections between blocks of neurons. The dashed arrows represent the modulatory circuits described above (i.e. the connections departing from the three internal neurons that reach and modulate the state of the nine visual sensory neurons in (c) or the state of the three internal neurons in (d)).

The free parameters of the robots' neural networks (i.e. the strength of the connections, the biases of internal and motor neurons, and the time constants of internal neurons) are evolved (Nolfi and Floreano, 2000) as described in the following section.

3.3 The evolutionary process

The initial population consists of 100 randomly generated genotypes that encode the connection weights, the biases and time constants of 100 corresponding neural controllers (each parameter is encoded by 8 bits and normalized in the range $[-5.0, +5.0]$ in the case of the connection weights and neurons' biases and in the range $[0.0, 1.0]$ in the case of neurons' time constants). The 20 best genotypes of each generation are allowed to reproduce by generating 5 copies each, with 2% of their bits replaced with a new randomly selected value. The evolutionary process lasts 2000 generations, i.e. the process of testing, selecting and reproducing robots is iterated 10 times. The evolutionary process is replicated 10 times for each experimental conditions (see section 4).

Each robot is allowed to "live" for 20 trials of 6.6min each (4000 time steps of 100ms). The position and the orientation of the robot, and the position of the food elements, is randomly initialized at the beginning of each trial by ensuring that the distance between objects and between the objects and the robot is greater than 2cm. The number of green food elements is reset to five at the beginning of each trial. When the robot collides with an obstacle (the wall or a food element) it loses all its current energy and remains stacked for the rest of the trial.

The fitness of the robot is computed by simply calculating the average level of energy of the robot during its entire lifetime. The fitness value thus varies in the range $[0.0, 1.0]$. The optimal value cannot be calculated due to the complexity of the scenario and the fact that the initial positions of the robot and of the food elements vary. However, it is certainly lower than 1.0 due to the time necessary to forage during the initial part of the trial and to the limited amount of available green substance.

In consideration of the characteristics of the environment and of the robots' internal "metabolic" processes, to maintain their energy level as high as possible the robots should develop and display several behaviours, presumably: (1) avoid colliding with obstacles, (2) explore the environment to find food sources, (3) approach a green object until the object is ingested, (4) approach the red object and remain near it until a sufficient quantity of red substance has been absorbed, (5) approach the blue object and remain near it until a sufficient quantity of blue substance has been absorbed, (6) avoid various types of objects when appropriate, (7) wait before ingesting green objects to avoid consuming all available green substance, (8) avoid wasting energy by producing unnecessary movements. Moreover, since part of the required behaviours are likely to be mutually exclusive, they should develop an ability to select the behaviour that are most appropriate to the current robot/environmental state.

4. Results

By analysing the results of several replications of the experiments we observed that in all experimental conditions the evolving robots develop and display multiple behaviours as well as an ability to select them on the basis of the robot/environmental circumstances so as to maintain their energy level relatively high for most of their lifetime. Overall these results confirm the validity of the model proposed by Seth (1988, 2012) in a more complex experimental scenario.

The comparison of the results obtained in different experimental conditions indicates that the availability of a modulatory circuit that facilitates the development of selective attention capacities lead to better results and to qualitatively different solutions in term of both behaviour realization and selection. These results support the hypothesis that selective attention facilitates/enables action selection and that the development of action selection strategies affect the development of

behaviour capacities and vice versa, i.e., the hypothesis that the mechanism enabling the development of a more effective action selection strategy might also enable the development of more effective behavioural capacities, and vice versa.

In section 4.1 we describe how performance varies quantitatively within the different experimental conditions. In section 4.2 we analyse the behaviour of evolved robots and how their behaviour varies in the different experimental conditions.

4.1 Comparative results obtained in different experimental conditions

The experiments have been replicated in the following four experimental conditions:

- A. The robots' neural network includes six internal neurons that receive connections from the camera and proprioceptors neurons and from themselves. The motor neurons receive connections from all the external sensory neurons and from the internal neurons (Figure 2a).
- B. The neural controller includes six internal neurons that receive connections from all the external and proprioceptors neurons and from themselves. The motor neurons receive connections from the internal neurons only (figure 2(b)). This experimental condition has been included to verify whether the lack of direct sensory-motor connections, that might encode direct sensory-motor reflexes, is beneficial or detrimental.
- C. The network includes three internal neurons that receive connections from the camera and the proprioceptors neurons and from themselves and three internal neurons that receive connections from the proprioceptors only. The second block of three internal neurons is used to set the gain of the corresponding three blocks of camera sensors (see Figure 2c). The motor neurons receive connections from the infrared sensors and from the first block of internal neurons. This experimental condition has been included to verify whether a neural modulatory circuit that regulates the activity of camera sensors can facilitate the development of an early selective attention capacity that in turn can facilitates/enables the development of an appropriate action-selection capacity. For a related connectionist model of selective attention that rely on top-down inhibitory mechanism please see Schrobsdorff et al. (2012).
- D. The network includes three internal neurons that receive connections from the camera and the proprioceptors neurons and from themselves and three internal neurons that receive connections from the proprioceptors only. The second block of three internal neurons is used to set the gain of the corresponding three internal neurons of the first block. The motor neurons receive connections from the infrared sensors and from the first block of internal neurons (Figure 2(d)). This experimental condition has been introduced to compare the adaptive advantage of regulatory mechanisms operating on early or successive levels of sensory processing (experimental condition C and D respectively).

It is important to notice that in principle evolving robots can solve their adaptive problem in any of the four experimental conditions. More specifically the presence of recurrent connections in the internal layers in all conditions could enable the robots to develop an effective action selection capability irrespectively from the presence/absence of top-down regulatory mechanisms.

The results show that condition C, i.e., the condition in which robots are provided with a modulatory circuit operating at the level of the visual sensory neurons, leads to better performance with respect to the other three conditions. The performance achieved in the other three conditions is rather similar overall, although robots' evolved in condition B display poorer fitness on the average (Figure 3, right).

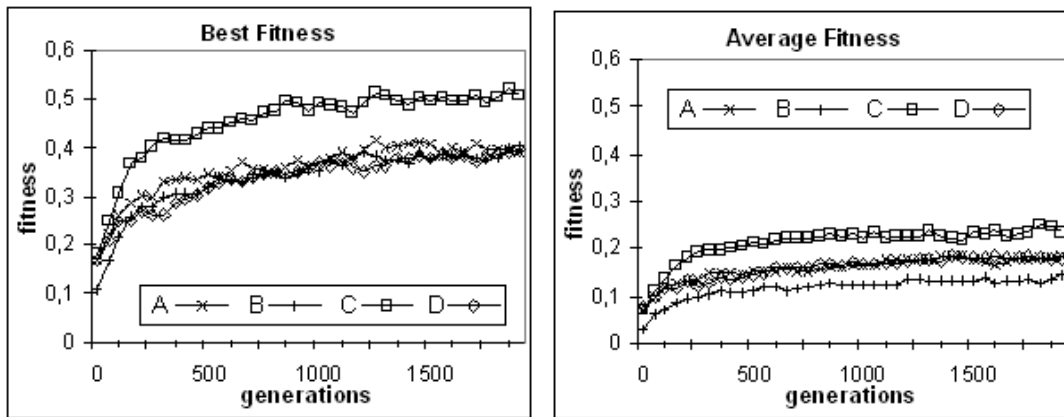


Figure 3. Fitness of the best individuals of each generation in the four experimental conditions. The letter A, B, C, and D refer to four corresponding neural architectures described in the text above. **Left:** Results for the best replication of each experimental condition. **Right:** Average result of 10 replications for each experimental condition

4.2 Behavioural analysis

To disentangle the robots' behavioural skills from the robots' action selection skills and to analyse the robots' ability to collect the food elements of different types, we post-evaluated evolved individuals in test conditions in which the need of substance was artificially manipulated. More specifically, we set the activation level of the proprioceptors neurons of the robots so as to reflect a situation in which one substance was completely missing while the other two substances were fully available and we analysed the average time required by the robots to ingest the missing substance (Figure 4). As shown in the Figure, performance does not vary too much between the four experimental conditions. More specifically, the robots evolved in condition C are the quickest in reaching green food elements but are slower than other robots in reaching the other two food elements. Overall, this analysis suggests that the higher performance of robots evolved in condition C is not due to a capacity to display more effective elementary foraging behaviours but rather to the ability to select among behaviours and/or to display better integrated behaviours.

The variation in performance observed between food elements of different types is probably due to the fact that finding and reaching a green objects is easier than reaching blue and red objects (due to the larger number of green objects available in the environment) and to the stronger adaptive pressure set on the foraging of green and blue food elements than of red elements (due to the fact that the availability of green and blue substances constitutes a pre-requisite for generating energy).

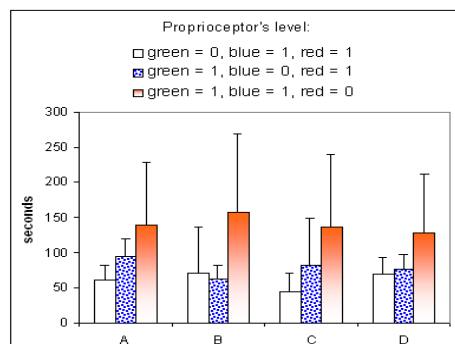


Figure 4. Average time required to reach green, or blue, or red food elements in a control condition in which the robots' lacked a selected substance only. The test is performed after 50s in which the robots are allowed to interact with the environment in normal conditions. Data averaged over 100 trials and over the best 10 evolved individuals of the 10 corresponding replications. The four groups of histograms display the data obtained in the four experimental conditions (A, B, C, and D).

Indeed, the analysis of the foraging actions carried out by evolved robots over time (Figure 5) and the analysis of the way in which evolved robots arbitrate between the different behaviours on the basis of their current internal state (Figure 7 and 8) indicate that the robots evolved in condition C display better arbitration strategies than robots evolved in the other experimental conditions.

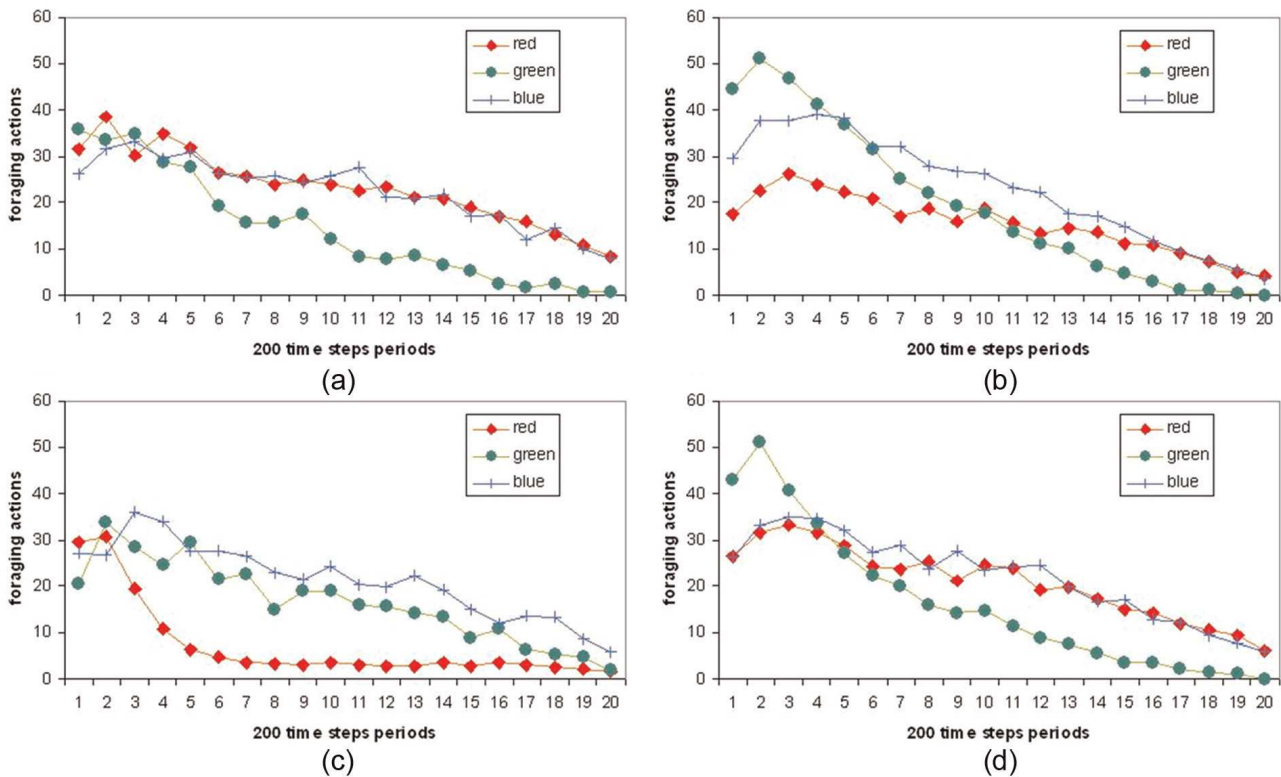


Figure 5. Average number of foraging actions displayed by robots evolved in experimental condition A, B, C, and D. Each curve shows the average number red, green, and blue foraging actions displayed during 20 successive periods of 200 time-steps forming trials of 4000 time-steps. Average results of the best 10 robots evolved in 10 corresponding replications of the experiment.

		20 successive periods of 200 time-steps																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Green	C vs A	✓										✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	C vs B	✓	✓	✓	✓	✓	✓	✓	✓	✓											
	C vs D	✓	✓	✓	✓	✓	✓	✓	✓	✓											
Red	C vs A			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	C vs B					✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	C vs D			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Blue	C vs A	✓										✓									✓
	C vs B	✓									✓										
	C vs D	✓									✓	✓									

Figure 6. Statistical differences between the green, red, and blue foraging actions displayed by the robots evolved in condition C versus the other three experimental conditions during each of the 20 successive time periods. Ticked box indicate the cases in which $p < 0.05$ according to a Mann–Whitney U test.

By analysing the number foraging actions we can see how they tend to decrease over time in all experimental conditions (see Figure 5). This can be explained by considering that green food elements tend to reduce as a result of the robots' foraging behaviour and by the fact that the robots might run out of energy and/or remain stacked for the rest of the trial. However, by analysing the number of green food elements collected over time we can see how the foraging actions decrease less markedly in condition C than in the other experimental conditions. In other words, C-robots manage to keep collecting green food elements for a longer period of time with respect to the other robots (Figure 5). This indicates that C-robots display a better capacity to postpone green-foraging actions to the phases in which the need of green substance is high thus reducing the risk to consume

too quickly the limited number of green food elements. Moreover, by analysing the number of time in which robots forage the red substances we can see how C-robots collect red food substance only during the first $\frac{1}{4}$ of the trial. This indicates that, unlike to the other robots, C-robots manage to successfully ignore the red food elements after a sufficient amount of red substance has been collected (see also below). As reported in Figure 6, the differences in the number red-foraging behaviours between condition C and the other three conditions are statistically significant from phase 6 to phase 16 in all phases. Similarly, the difference in the number of green-foraging behaviours between condition C and the other three conditions are statistically significant from phase 12 to phase 19 during most of the phases.

To analyse in more details how robots select actions we segmented the behaviour produced by robots into a sequence of functionally different behaviours (reach food X, avoid food X, and wandering behaviours) and we analysed the relation between the robots' internal state and the robots' behaviour. More specifically, sequences of sensory-motor interactions in which a robot visually perceives a given food element and acts so as to continue to perceive the same food element until the robot finally reaches it are categorized as "reach food X behaviours". Sequences of sensory-motor interactions in which the robot visually perceives a food X and act so as to stop perceiving the food after some time without reaching it are categorized as "ignore food X behaviours".

Figure 7 and 8 display the probability that each of the six reach and avoid food behaviours is exhibited (i.e. selected) for each possible combination of green and blue internal states. As can be seen, the analysis conducted on the best robot evolved in condition C (see Figure 7), indicates that it displays: a reach-blue behaviour in most of the circumstances in which the need of the blue substance exceed 0.25; a reach-green behaviour when the need of green substance is high (i.e. $>$ than 0.5); and a reach-red behaviour during the initial part of the trial when the need of all substances is maximum. The high probability of executing any of the three foraging behaviour during the initial part of the trials in which the need of all substances is maximum can be explained by considering that in this phase the exhibition of any foraging behaviour is appropriate.

The same analysis conducted on the best robots evolved in the other experimental conditions shows in general that these individuals are less capable to prioritize behaviours on the basis of their internal needs. For reasons of space we only display the data for the best robot evolved in the experimental condition A (Figure 8) but the overall pattern is similar for conditions B and D. As can be seen, A-robots tend to erroneously exhibit a reach-red behaviour (rather than an ignore-red behaviour) also when the need of green substance is high. Moreover they erroneously display an ignore-green and an ignore-blue behaviours also in circumstances in which the need of the two corresponding substances is high (i.e. near or over 0.5).

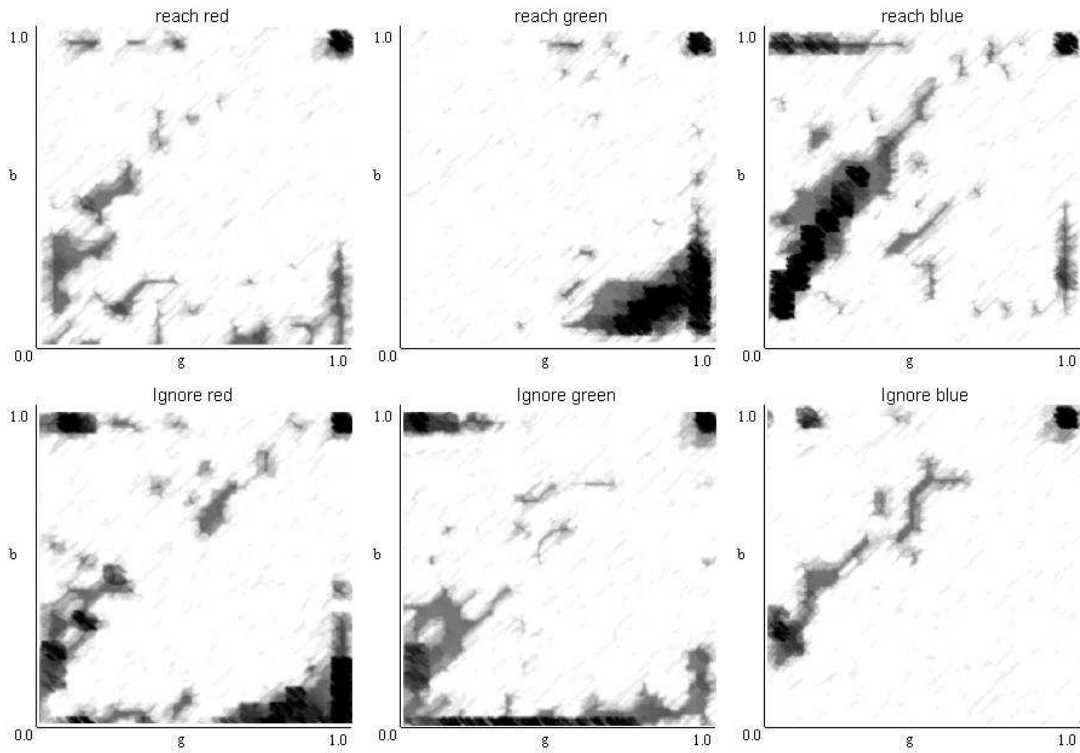


Figure 7: Probability of execution of the three reach-x and ignore-x behaviours in each possible combination of blue/green internal state. Data computed by partitioning the internal states into 21 intervals equally distributed between 0.0 and 1.0. In each picture the vertical and horizontal axis indicate the need of blue and green substances (g, b). The darkness of the colour indicate the probability that the corresponding behaviour is displayed for each possible combination of internal states normalized between 0 (white) and 100% (black). Analysis performed on the best evolved individual of experimental condition C for 100 trials.

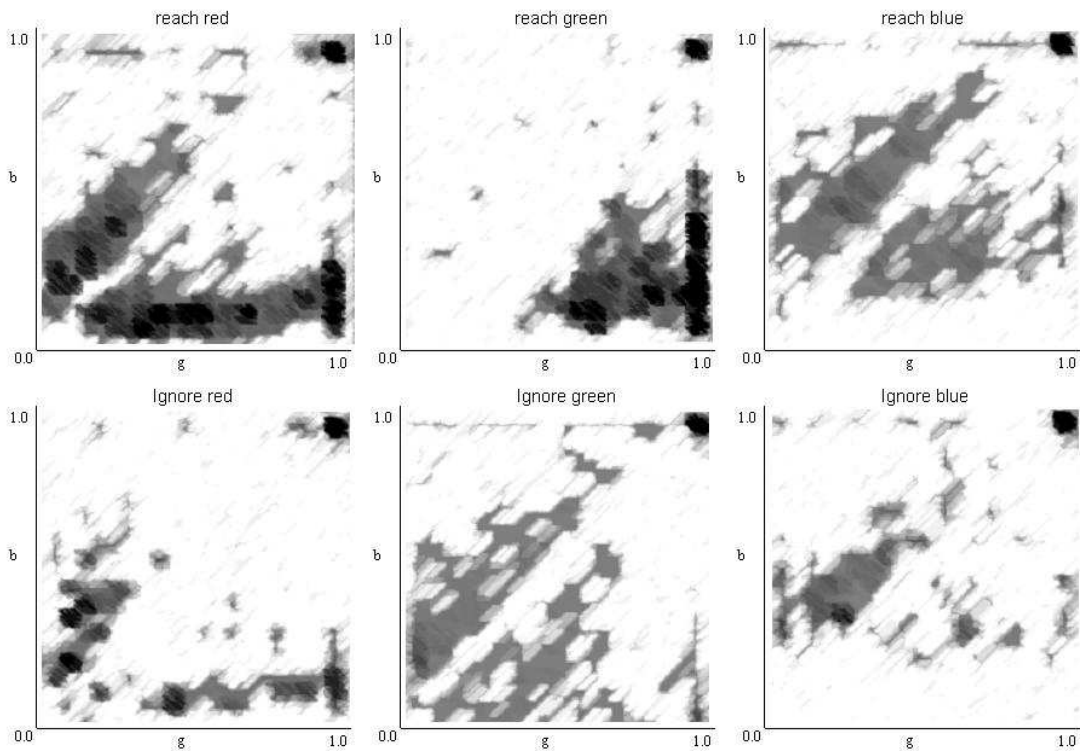


Figure 8: Probability of execution of the three reach-x and ignore-x behaviours in each possible combination of blue/green internal state (see caption of Figure 5). Analysis performed on the best evolved individual of experimental condition A for 100 trials.

In addition to a better prioritization capacity, the best robot evolved in experimental condition C is characterized by more integrated behaviours and by a capacity to effectively regulate its action selection strategy over time.

The development of better integrated behavioural capability can be appreciated by considering that the reach-blue behaviour in this robot is realized by approaching the blue object and then by keep circling around it, even when there is no more need to collect blue substance, until the need of green substance reach a given threshold. The reach-green behaviour is realized by approaching the green object and then by making a u-turn of about 180 degrees that allows the robot to face back toward the area from which it reached the green object. The alternation of these two behaviours allows the robot to move back and forth between the blue and green food elements without the need to explore the environment to find the blue food. Indeed, the fact that the robot prioritizes the foraging of the blue substance with respect to the green substance and the fact that it keeps circling around the blue food until it selects the reach-green behaviour implies that it systematically travels from the blue food element to one of the green food element. This, in turn, ensures that by moving toward a green food element from the blue area and by then making a u-turn the robot will orient back towards the blue food element. The robots evolved in the other experimental conditions do not shows this type of integrated behaviour (i.e. do not show the keep-circling behaviour around the blue food element and do not travel directly from the blue to the green food elements). Consequently, they cannot exploit the benefits associated with the capacities to keep the level of the ingested blue substance high (the level of the blue substance only decreases during the travels toward and backward the green food elements) and to travel directly toward the blue area even by distance when it can hardly be perceived.

The integrated behaviours shown by the best robot evolved in condition C explains the diagonal stripe of frequencies associated to the exhibition of the reach-blue behaviour shown in Figure 7. Indeed, the exhibition of the integrated behaviours described above implies that the need of blue and green substances displayed when the robot exhibits a reach blue behaviour correlate linearly with the amount of blue and green food substance consumed by the robot during the travel toward the green food and then back toward the blue food (which in turn depends on the distance between the blue food and the currently selected green food).

The capability of the best C-robot to effectively regulate the action selection strategy over time can be appreciated by analysing how the state of the regulatory neurons (that determine the gain of the red, green, and blue visual photoreceptors) varies during the first 30s of a trial in which the robot is situated in an environment which does not contain any food element (see Figure 9). Remember that at the beginning of a trial the robot lacks all the three substances.

As shown in the Figure, the robot initially pays attention to the red and blue objects but not to the green objects. The attention to red food can be explained by considering that it should be ingested as soon as possible to maximize the amount of energy than can be later be extracted from geen and blue substances. The higher attention toward blue than green food can be explained by the fact that the blue food is more difficult to find and by the fact that the robot is able to navigate efficiently from blue to green food and back thanks to the integrated behaviours described above. The most interesting aspect however is constituted by the fact that the attention toward the red and green substances decreases and increase respectively over time during the first 23s of the trial. Increasing the attention toward the green food over time, in fact, enables the robot to reduce the risk of running out of energy in the attempt to first gather the red substance.

As for the integrated behavioural capacity described above, this capacity to regulate the action selection process over time during phases in which the robot lacks all the three substances is observed only in some of the robots evolved in the experimental condition C.

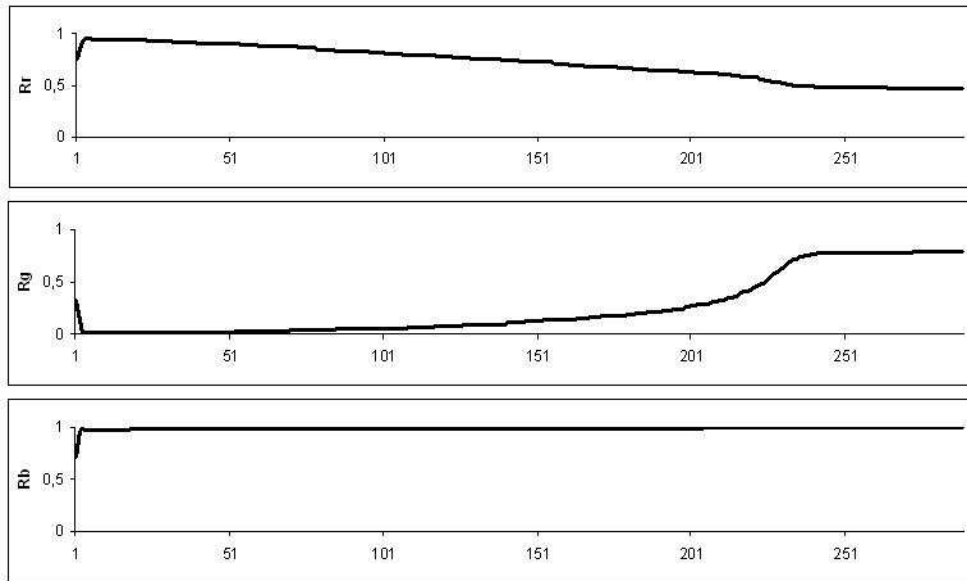


Figure 9: Activation of the regulatory neurons (Rr, Rg, and Rb) that determine the gain of the corresponding blocks of red, green, and blue photoreceptors during the first 30s of a trial in which the robot is situated in an food deprived environment.

4. Discussion and conclusion

We presented a series of experiments in which embodied and situated agents are evolved for the ability to collect different food substances to maximize the energy that they can generate through their internal metabolic process. The characteristics of the robots, of the environment, and of the task create an adaptive pressure toward the development of several behavioural capabilities as well as toward the development of an ability to select among mutually exclusive behaviours. More precisely, the experimental setup creates the adaptive condition for the co-development of integrated behavioural and action selection capabilities.

The fact that this method does not require to hand-craft the agents behavioural capabilities (i.e. the number of behavioural skills, the way in which behaviours are realized, and the control mechanisms that enable behaviour exhibition) allows us to avoid the risk to introduce in our model unnecessary complications and biases caused by the external perspective of the designer that might affect the need of specific action selection mechanisms and/or the way in which action selection could be realized. More specifically, this method allows us to build a model that enables the exploitation of behavioural and action selection capabilities that emerge from the agent/environmental interactions. Moreover, it allows us to avoid the misleading assumption that the mechanism underlying the generation of behaviours consists of internal correlates (e.g. control modules or layer) or the behaviours themselves and that action selection is realized by arbitrating between such internal units (an assumption that is often implicit in many models of action selection as discussed by Seth [2012]).

Overall the obtained results confirm that, as illustrated by Seth (1988, 2012), a capacity to display multiple behavioural skill and to select among them (i.e. by prioritizing action selection on the basis of the agents internal needs, by allowing continuous behavioural sequences to be strung together, by exhibiting opportunism etc., Werner [1994]) can arise from the concurrent activity of multiple sensory-motor processes without postulating any dedicated internal arbitration mechanism.

However, the comparison of the results obtained in the experimental conditions in which the robots were or were not provided with regulatory neural mechanism that facilitates the emergence of covert selective attention demonstrates how it enables the development of more effective action and action selection capabilities. Our results thus provide additional evidences in support of the theories that claim that one key function of selective attention is indeed to control action (Newman,

1987; Tipper et al, 1988; Castiello, 1999). More specifically, our finding support the affordance competition theory (Cisek, 2012) that claims that sensory stimuli tend to evoke directly the actions afforded by them and that competing or interfering evoked actions are eliminated thanks to selective attention mechanisms that reduce the amount of information that is transformed into action-related representation. The action control role of selective attention also suggests that one of the primary functions of internal neural connections from the motor areas to the sensory areas observed in many species (Webb, 2004) could be that to facilitate the selection among competing actions.

The fact that C-robots, in which the regulatory mechanism operates at the level of the sensory neurons, outperform D-robots, in which the regulatory mechanism operates at the level of the internal neurons, suggests that the reduction of the information to be processed should occur as early as possible. However, further experiments should be carried out to verify the role of the other differences between the two experimental conditions.

Our obtained results also demonstrate the strict interdependence between action and action selection capabilities. Indeed, the comparison of the best robot provided with the neural modulatory circuit (experimental condition C) with the best robots obtained in the other experimental conditions demonstrated how the former robot outperforms the other robots not only with respect to its action selection capability but also with respect to the way in which actions are realized and integrated. This result indicates that the mechanisms enabling the development of better action selection capabilities also create the condition for the development of better action capabilities and (possibly) vice versa.

Acknowledgments

This research has been supported by CNR under the European Science Foundation project Hierarchical Heterogeneous SWARM (H2Swarm).

References

- Bassler U. and Buschges A. (1998). Pattern generation for stick insect walking movements. Multisensory control of a locomotor program. *Brain Research Reviews*, 27: 65-88.
- Bonfiglioli C. and Castiello, U. (1998) Dissociation of covert and overt spatial attention during prehension movements: selective interference effects *Percept. Psychophys.* 60: 1426–1440.
- Brooks R.A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2:14-23.
- Bryson J.J. (2000). Hierarchy and sequence versus full parallelism in action selection architectures. In J-A. Meyer, A. Berthoz, D. Floreano, H. Roitblat, & S. W. Wilson (Eds.), *From Animals to Animats 6: Proceedings of the Sixth International Conference on Simulation of Adaptive Behaviour*. Cambridge MA: MIT Press.
- Buschges A. and Wolf H. (1999). Phase dependent presynaptic modulation of mechanosensory signals in the locust flight system. *Journal of Neurophysiology* 81: 959-962
- Castiello U. (1996) Grasping a fruit: selection for action *J. Exp. Psychol. Hum. Percept. Perform.* 22: 582–603
- Castiello U. (1999). Mechanisms of selection for the control of hand action. *Trends in Cognitive Sciences*, 3 (7): 264-271.
- Chieffi, S. et al. (1993) Study of selective reaching and grasping in a patient with unilateral parietal lesion: dissociated effects of residual spatial neglect. *Brain* 116:1119-1137.
- Cisek P. (2012). Cortical mechanism of action selection: the affordance hypothesis. In A. Seth, T.J. Prescott, and J.J. Brysonn (Eds.). *Modelling Natural Action Selection*. Cambridge, U.K.: Cambridge University Press.

- Duncan-Johnson C.C. and Koppell B.S. (1981). The Stroop effect: brain potentials localize the source of interference. *Science* 214:938-940.
- El Manira A, DiCaprio R.A, Cattaert, D., and Clarac, F. (1991). Monosynaptic interjoint reflexes and their central modulation during fictive locomotion in crayfish. *European Journal of Neuroscience* 3:1219-1231.
- Gebhart M. and Honnegger H.W. (2001). Physiological characterisation of antennal mechanosensory descending interneurons in an insect (*Gryllus bimaculatus*, *Gryllus campestris*) brain. *Journal of Experimental Biology*, 204: 2265-2275.
- Gibson J.J. (1979). *The Ecological Approach to Visual Perception*. Boston MA: Houghton Mifflin.
- Goldberg M.E. and Segraves M.A. (1987). Visuo-spatial and motor attention in the monkey. *Neuropsychologia*, 25:107-118.
- Howard L.A and Tipper S.P. (1997) Hand deviations away from visual cues: indirect evidence for inhibition *Exp. Brain Res*, 113: 144–152.
- Jackson S.R., Jackson G.M. and Rosicky, J. (1995) Are non-relevant objects represented in working memory? The effect of non-target objects on reach and grasp kinematics *Exp. Brain Res*, 102:519-530.
- Maes, P.(1990). Situated agents can have goals. *Robotics and Autonomous Systems*, 6, 49-70.
- Miller J. and Hackley S.A. (1992). Electrophysiological evidence for temporal overlap among contingent mental processes. *J. Exp. Psychol. Gen*, 121:195-209.
- Mondada, F., Bonani, M., Raemy, X., Pugh, J., Cianci, C., Klaptocz, A., Magnenat, S., Zufferey, J.-C., Floreano, D. and Martinoli, A. (2009). The e-puck, a Robot Designed for Education in Engineering. *Proceedings of the 9th Conference on Autonomous Robot Systems and Competitions*, 1(1) pp. 59-65.
- 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, A. Berthoz, D. Floreano, H. Roitblat, & S. W. Wilson (Eds.), *From Animals to Animats 6: Proceedings of the Sixth International Conference on Simulation of Adaptive Behaviour*. Cambridge MA: MIT Press.
- Neumann, O. (1987) Beyond capacity: a functional view of attention. In H. Heuer, and A.F. Sanders (Eds.), *Perspectives on Attention and Action*. Erlbaum Press.
- Nolfi S. & Floreano D. (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. Cambridge, MA: MIT Press/Bradford Books.
- Nolfi S. & Gigliotta O. (2010). Evorobot*: A tool for running experiments on the evolution of communication. In S. Nolfi & M. Mirolli (Eds.), *Evolution of Communication and Language in Embodied Agents*. Berlin: Springer Verlag.
- Parisi D., Petrosino G. (2010). Robots that have emotions, *Adaptive Behavior*, vol. 18(6), pp. 453-469.
- Powers W. T. (1973). Feedback: Beyond behaviourism. *Science*, 179 (4071): 351-356.
- Prescott T. J. (2008). Action selection. *Scholarpedia*, 3(2):2705.
- Prescott, T. J. (2007). Forced moves or good tricks in design space? Landmarks in the evolution of neural mechanisms for action selection. *Adaptive Behavior*, 15(1), 9-31.
- Riddoch M.J., Humphreys G.W. and Edwards M.G. (2000). Visual affordances and object selection. In S. Monsell and J. Driver (Eds.). *Attention and Performance XVIII*. Cambridge, MA: MIT Press.
- Ruini F., Petrosino G., Saglimbeni F., Parisi D. (2010). The strategic level and the tactical level of behaviour, in J. Gray and S. Nefti-Meziani (eds.), *Advances in cognitive systems*, Herts, IET Publisher.
- Seth A. (1988). Evolving action selection and selective attention without action, selection, and attention. In R. Pfeifer, B. Blumberg, J.A. Meyer, and S. Wilson (Eds.) *Proceeding of the Fifth International Conference on the Simulation of Adaptive Behaviour*. Cambridge, MA: MIT Press.

- Seth A. (2012). Optimized agent based modelling of action selection. In A. Seth, T.J. Prescott, and J.J. Brysonn (2012). *Modelling Natural Action Selection*. Cambridge, U.K.: Cambridge University Press.
- Seth A., Prescott T.J., and Brysonn J.J. (2012). *Modelling Natural Action Selection*. Cambridge, U.K.: Cambridge University Press.
- Schrobsdorff H, Ihrke M, Behrendt J., Hasselhom M. and Hermann J.M. (2012). Inhibition in the dynamics of selective attention: an integrative model for negative priming. *Frontiers in Psychology*, 3 (491): 1-21.
- Tipper S.P., Howard A. and Houghton G. (1998). Action-based mechanism of attention. *Philosophical Transactions of the Royal Society of London B*, 353(1373): 1385-1393.
- Tipper S.P., Howard L.A. and Houghton G. (1998) Action-based mechanisms of attention *Philos. Trans. R. Soc. London Ser. B*: 353:1385-1393.
- Tipper S.P., Howard L.A. and Jackson, S.R. (1997) Selective reaching to grasp: evidence for distractor interference effects. *Visual Cognition*, 4: 1-38.
- Webb B. (2004). Neural mechanisms for prediction: do insects have forward models? *Trends in Neurosciences*, 27 (5): 278-282.
- Tyrrel T. (1993). The use of hierarchies for action selection. *Adaptive Behaviour*, 1: 387-420.
- Werner G. (1994). Using second-order neural connections for motivation of behavioural choice. In D. Cliff, P. Husband, J.A. Meyer, and S. Wilson (Eds). *From Animals to Animats 3: Proceeding of the Third International Conference on the Simulation of Adaptive Behaviour*. Cambridge, MA: MIT Press.