Theory Biosci. (2016) 135:201–216
DOI 10.1007/s12064-016-0233-y

ORIGINAL PAPER

Behavioural plasticity in evolving robots

Jônata Tyska Carvalho[1,2] · Stefano Nolfi[1]

Received: 10 September 2015 / Accepted: 9 July 2016 / Published online: 21 July 2016

Abstract In this paper, we show how the development of plastic behaviours, i.e., behaviour displaying a modular organisation characterised by behavioural subunits that are alternated in a context-dependent manner, can enable evolving robots to solve their adaptive task more efficiently also when it does not require the accomplishment of multiple conflicting functions. The comparison of the results obtained in different experimental conditions indicates that the most important prerequisites for the evolution of behavioural plasticity are: the possibility to generate and perceive affordances (i.e., opportunities for behaviour execution), the possibility to rely on flexible regulatory processes that exploit both external and internal cues, and the possibility to realise smooth and effective transitions between behaviours.

Keywords Behavioural plasticity · Evolutionary robotics · Multiple behaviours · Autonomous robotics · Modularity · Action switching

✉ Jônata Tyska Carvalho
jonatatyska@gmail.com

Stefano Nolfi
stefano.nolfi@istc.cnr.it

[1] Institute of Cognitive Sciences and Technologies, National Research Council (CNR), Via S. Martino della Battaglia, 44, 00185 Rome, Italy

[2] Center for Computational Sciences (C3), Federal University of Rio Grande (FURG), Av. Italia, km 8, Rio Grande 96203-900, Brazil

Introduction

Behavioural plasticity is a special case of plasticity—“the ability of an organism to react to internal or external environmental inputs with a change in form, state, movement, or rate of activity” (West-Eberhard 2003, p. 33). It involves the capability to display multiple behavioural responses, which might differ in a continuous or discontinuous way, in a condition-sensitive manner (Komers 1997).

Behavioural plasticity constitutes a key aspect of animal behaviour. Indeed, behaviours are often organised in functionally specialised subunits governed by switch and decision points (Gallistel 1980). Examples of elaborate behaviours including several different phases regulated through a rich set of context-dependent rules include the courtship behaviour of the grasshopper (Otte 1972), the reproduction behaviour of female canaries (Hinde 1970), web construction and predation behaviours in spiders (Eberhard 1988; Jackson and Wilcox 1993).

Behavioural plasticity is essential for enabling organisms to adapt to variations of their external and/or internal environment. In that respect, it is important to consider that what matters, from the point of view of the adapting individuals, is the organism’s perceptual environment (i.e., the characteristics of the environment that the organism perceives given its sensory system and its relative location in the environment). This means that all environments are variable, from the perspective of an organism that is situated and performs actions in an environment, independently of whether they appear variable or not from the perspective of an external observer.

In this paper, we analyse experimentally how evolving robots can acquire and display behavioural plasticity, i.e., a series of behaviours that are exhibited in a context-dependent manner. In particular, we analyse whether behavioural plasticity evolves during the course of the evolutionary process, which are the prerequisites for its evolution, and which are the mechanisms through which it is realised. The comparison of the results obtained in different experimental conditions indicates that the most important prerequisites for the evolution of behavioural plasticity are the possibility to generate and perceive affordances (i.e., opportunities for behaviour execution), and the possibility to regulate, in a flexible manner, the alternation of the different sub-behaviours and the transitions between sub-behaviours.

Behaviour, multiple behaviours and behavioural plasticity

For the sake of clarity, it is important to specify what we mean by behaviour, multiple behaviours and behavioural plasticity. In the context of agents that are embodied and situated, *behaviour* is the dynamical process that originates from agent/environmental interactions. At any time step, the environment and the agent/environment relation co-determine the body and the motor reaction of the agent that, in turn, co-determine how the agent/environment relation and/or the environment vary. Sequences of interactions lead to a dynamical process that extends for a certain period of time: the agent's behaviour.

We use the term *overall behaviour* to indicate the entire behaviour displayed by an agent, i.e., the behaviour displayed by an organism during its entire lifetime. Moreover, we use the term function/s to indicate the adaptive role of behaviour, e.g., the overall behaviour displayed by an organism can have the function of enabling the organism to survive and reproduce.

Behaviour might be characterised by a modular organisation with somewhat semi-discrete and semi-dissociable subunits (West-Eberhard 2003), or *sub-behaviours,* playing different functions (or sub-functions). When sub-behaviours display a modular organisation as well, the behaviour displays a hierarchical organisation characterised by multiple-levels (e.g., lower-level behaviours, higher-level behaviours, overall behaviour, see Nolfi 2009). We used the term semi-discrete and semi-dissociable to emphasise the fact that conceptualising sub-behaviours as a collection of independent subunits is misleading, since sub-behaviours are only partially independent from each other. The modular organisation of behaviour, therefore, is characterised by both discreteness and evidence of boundaries between sub-behaviours and by connectedness and integration among them (West-Eberhard 2003). After all, even individual organisms are not completely independent units, given that they also show a significant level of connectedness and interdependence with conspecifics, in most of the species. Notice that the modular organisation of behaviour should not be confused with the modular organisation of the agent's nervous system.

The term *multiple behaviours* refers to behaviours characterised by a modular organisation, i.e., characterised by the presence of multiple semi-independent sub-behaviours. In behaviours displaying several levels of organisation, the presence of multiple semi-independent behavioural units characterises all levels of organisation, except the level of the overall behaviour. As an example, we can consider the behaviour of a tennis player during a game that can be divided in a series of semi-independent sub-behaviours such as serve and volley (in which the player serves and then charges forward to the net), lob (a shot in which the ball is lifted high above the net) etc.

The term *behavioural plasticity* refers to agents displaying behaviours characterised by a modular organisation and displaying the capability to regulate the exhibition of the different sub-behaviours on the basis of their internal and external environment. In the example of the tennis player, behavioural plasticity refers to the capability of displaying multiple behaviours such as these described above and to the capability to select the appropriate behaviour depending on the game context, for example the ability to execute a drop shot behaviour, that consists in hitting the ball just over the net, when the opponent is far from it. The term behavioural plasticity should not be confused with neural plasticity, e.g., fine-grained modifications of the connection weights of the agent's nervous system (see Nolfi and Floreano 1999).

Whether behavioural units or sub-behaviours should be considered as real entities eligible for scientific analysis or subjective entities that only exist in the eyes of the observer represents an open question. Indeed, although many biologists assume that behaviour is organised in semi-discrete units with specialised functions (Mitchell 1990; Barlow 1977; Gallistel 1980; Wenzel 1993; West-Eberhard 2003), others consider behavioural units as useful fictions at best (Fentress 1983). Within the Artificial Life and Robotics community, the notion of behavioural unit has a relatively clear and non-controversial meaning in the context of behaviour-based architectures (Brooks 1986) in which different modules or layers are responsible for the production of alternative corresponding sub-behaviours (i.e., in a situation in which there is a one-to-one correspondence between behavioural units and agent's control modules and in which the control modules are separated by clear boundaries). Whether robots operating on the basis of non-modular neural controllers can properly make use of multiple behaviours, as well, represents an open question (see Tani and Ito 2007; Prescott 2008; Nolfi 2009). The attempt to resolve this issue is outside the scope of this paper. For intellectual honesty, we clarify that together with several authors cited above, we assume that the behaviour of an agent can have a modular organisation even when the behavioural units do not correspond to clearly identifiable components of the agent. As argued by West-Eberhard (2003, p. 63), we believe that "it would be foolish to deny the modular properties of phenotypic organization just because there are connections and indistinct borders around the subunits we recognize as trait. There can be no doubt that there exists behavioural subroutines or subunits, for they are distinguishable from others in form, function, and discreteness, and sometimes in gene expression …". Moreover, we assume that the presence or the lack of a modular behavioural organisation can have important consequences (e.g., on agents' performance and on agents' ability to develop new skills).

Relation to the state of the art

Evolutionary robotics (Nolfi and Floreano 2000; Nolfi et al. 2016) concerns the synthesis of population of embodied and situated robots that develop their skills autonomously as a result of an evolutionary process based on selective reproduction and variation. In this context, the study of behavioural plasticity has been addressed indirectly in the following three research lines.

The first research area concerns the study of the combination of evolution and learning (Nolfi and Floreano 1999). Nolfi and Parisi (1997), in particular, showed how evolving robots manage to successfully vary their behaviour during the course of their life to adapt to variations of objects reflectance. Floreano and Nolfi (1997) showed how evolving predator robots vary their predation strategy on the basis of the behaviour displayed by the escaping prey so as to successfully capture it.

The second line of research addresses the study of the potential advantage of evolutionary algorithms supporting the evolution of modular neural controllers. The rationale behind this is that the availability of separated neural modules can facilitate the exhibition of behaviours characterised by a modular organisation. In some cases, this objective was realised by providing the neural controllers with a varying number of neural modules arbitrated on the basis of a co-evolved arbitration mechanism (Calabretta et al. 2000; Schrumand and Miikkulainen 2012). In other studies, instead, it was realised by genetically encoding the connectivity between the neurons, i.e., by enabling the evolutionary process to select architectures displaying clusters of neurons with many intra-connections and few inter-connections (Bangard 2011; Verbancsics and Stanley 2011; Huizinga et al. 2014).

Finally, the third line of research concerns the study of action selection (behaviour selection for consistency with the terminology we are using), i.e., the capacity to select between alternative behaviours afforded by the current organism/environmental context (Seth et al. 2012). In most of the cases, evolutionary studies conducted in this area concern the evolution of an ability to arbitrate hand-designed control modules producing predetermined behaviours (e.g., Gonzales et al. 2000; Rahim et al. 2014). In other cases, however, the behaviours were evolved as well (Izquierdo and Bührmann 2008; Seth 2012; Petrosino et al. 2013; Williams and Beer 2013). In these experiments, however, the synthesis and the exhibition of multiple behaviours represented the only possible viable solution since the evolving robots were required to carry on mutually exclusive tasks [e.g., eating or avoid eating a specific food type (Seth 2012; Petrosino et al. 2013) or moving on the basis of a wheeled or legged actuators (Williams and Beer 2013)].

In this paper, we run a series of experiments that aim to study whether behavioural plastic solutions evolve, whether they provide advantages with respect to non-plastic solutions and which are the factors that represent necessary prerequisites for the evolution of behavioural plasticity. As we will see, our results indicate that behavioural plastic solutions can evolve also when the adaptive task does not require the accomplishment of multiple conflicting functions. Moreover, our results indicate that behavioural plastic solutions might enable the evolving agents to achieve higher performance. The analysis of our experiments indicates that the most important prerequisite for the evolution of behavioural plasticity is constituted by the capability to perceive and generate affordances, i.e., opportunities for behaviours (Gibson 1979; Chemero 2011). This capability depends on the richness of the robot's perceptual environment that, in turn, depends on the richness of the robot's internal and external environments, on the richness of the robot's sensory–motor system, and on the ability to exploit sensory–motor coordination. Moreover, the analysis indicates the importance of using flexible regulation mechanisms that rely on both external and internal cues. Finally, the obtained results demonstrate the importance of the connectedness between sub-behaviour and the importance of providing the agents with mechanisms that enable them to realise a smooth and effective transition between sub-behaviours.

The method

To study this issue, we decided to consider a cleaning experimental scenario in which a wheeled robot need to vacuum-clean the floor of an unknown in-door environment. We choose this problem since it represents the first (and still the more significant) successful application domain of autonomous robot solutions (Roomba, the first autonomous vacuum-cleaning robot developed by iRobots® under the supervision of Rodney Brooks and commercialised from 2002 has been sold in more than 10 million units to date, see iRobot 2013). Rather than designing the controller by hand, we studied whether effective controllers can be developed from scratch through an evolutionary method in which the evolving robots are selected on the basis of the percentage of successfully cleaned surface, i.e., on the basis of a scalar value that rates their overall ability to perform the task.

It is important to point out that we choose this domain also because it involves the execution of a task with a single goal (cleaning the environment) that does not necessarily require behavioural plastic solutions. This enables us to study whether and how behavioural plastic solutions evolve, whether and why they provide an advantage with respect to non-plastic solutions, and eventually which are the characteristics and functions of the evolved sub-behaviours. Domains involving multiple conflicting goals, such as those used in the literature addressing the study of action selection cited above, in fact, necessarily require the development of solutions characterised by multiple behaviours and, implicitly, constrain the number and type of required sub-behaviours.

The investigation of the cleaning problem also permits to compare our evolved solutions with those developed by companies that sell cleaning robots. In that respect, the fact that the behavioural policies displayed by different versions of the Roomba and by similar robots produced by other companies significantly differ (Ackerman 2010) demonstrates that finding the optimal solution/s of this problem is far from trivial.

The task, the environment and the robot

To evolve robots that are robust with respect to environmental variations, we evaluated each robot for three trials or cleaning sessions. At the beginning of each trial, the initial position and orientation of the robot in the environment, and the specific characteristics of the environment, like dimensions and object positions, in which it was situated in were randomly varied within limits.

Each trial lasted 6 min and 15 s. Although performing a precise comparison with the time required by commercial robot to clean completely or almost completely a surface with similar properties is impossible due to the lack of data (for some indications see Ackerman 2010), this represents a rather short period of time.

To compute the cleaning performance, we calculated the percentage of 20×20 cm non-overlapping areas visited by the robot at least once during a trial.

The experiments have been repeated in two different types of environments. In the first set of experiments, we used a concave environment (Fig. 1, left) constituted by a large central area and by four peripheral corridors that represent a room-like environment. The average environment had a central area with a size of 6.8 m^2 and four corridors with a size of 3.78 m^2 in total. The exact size of the environment, however, was randomly set at the beginning of each trial. This was realised by varying the height and width of the central area and of corridors of ± 33 and ± 18 %, respectively, during different trials. In the second set of experiment, we used a convex environment (Fig. 1, right) constituted by a rectangular room-like area including furniture. The rectangular area has a size of 12.2 m^2 $\pm$ 33 % and includes: a first rectangular object with an area of 0.93 m^2 $\pm$ 10 %, a second rectangular object with an area of 0.17 m^2, the legs of a table, and the legs of chairs (the number of chairs was randomly varied in the range [0, 4]). The x and y coordinates of all the objects located over the plane were also varied during each trial within limits that prevented physical overlap.

The robot used was a MarXbot (Bonani et al. 2010), a differential drive wheeled robot with a diameter of 17 cm. The robot is equipped with 24 infrared sensors evenly distributed along the robot's body and capable of detecting

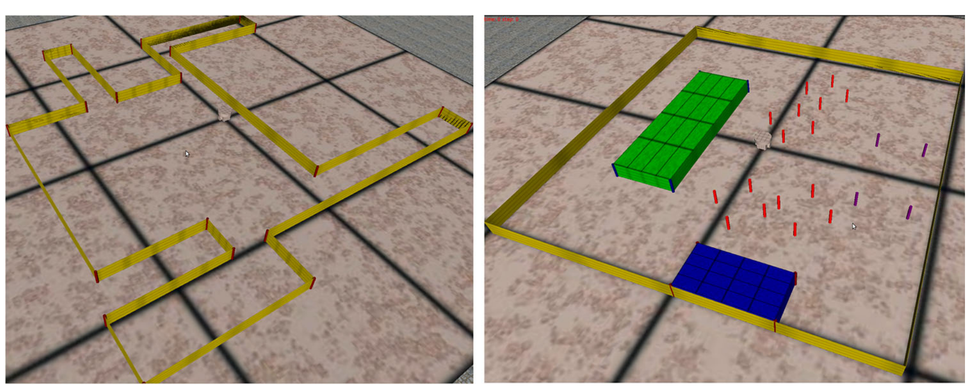

Fig. 1 Examples of concave and convex environments, *left* and *right*, respectively

objects in a range of 10 cm. Moreover, it is equipped with a rotating laser sensor capable of detecting obstacles at longer distance. Experiments were run in simulation using the FARSA open-software tool (Massera et al. 2013) that includes an accurate simulator of the robot and of the environment.

The robots' neural controller

The robots are provided with a neural network controller. In all experiments, the robots are equipped with eight sensory neurons that encode the average activation state of eight groups of three adjacent infrared sensors each and two motor neurons that encode the desired speed of the two robot's wheels. The sensory neurons are fully connected with the motor neurons and to hidden neurons (if present), and the hidden neurons are fully connected to the motor neurons. Hidden and motor neurons are provided with biases. The state of the hidden and motor neurons is computed on the basis of the logistic function. The state of the sensory neurons and the desired speed of the robot's wheels are updated every 50 ms. Experiments have been replicated in the following four experimental conditions:

(S) Simple: The robots are only provided with the infrared sensors

(R) Range sensor: The robots are provided with an additional sensory neuron that encodes the average distance of obstacles located within 1 m detected through the rotating laser range sensor. This sensor has been added to enable the robot to vary its behaviour in narrow versus open areas

(T) Time: The robots are provided with an additional sensory neuron that encodes the time passed since the beginning of the current cleaning session (trial), i.e., whose activation state linearly varies between 1.0 and 0.0 during the course of the trial. This sensor has been added to enable the robot to vary the behaviour during the course of cleaning sessions. Notice that this sensor enables the robot to access information extracted from the robot's internal environment (e.g., a robot clock situated inside the robot body), while the other sensors enable the robot to access information extracted from the external environment

(M) Modular: The neural controller is formed by three modules (each provided with eight infrared sensors connected to the two motor neurons) that are used during three subsequent phases of the trial of equal length. This modular neural controller was used to enable the robot to freely differentiate its behaviour during the three successive phases of the trial

To investigate whether the addition of internal neurons could enable the robot to achieve better performance, we carried out a second series of experiments in which the robot was also provided with an additional layer with three hidden neurons that received connections from all sensory neurons and projected connections to all motor neurons.

The connection weights and biases, that determine the robots' behaviour, are initially set randomly and evolved as described in the section below. The tool used to run the experiment can be downloaded from https://sourceforge.net/projects/farsa/. The source of the plugin that enables to replicate this experiment can be downloaded from http://sourceforge.net/p/farsa/code/HEAD/tree/farsaPlugins/cleaningExperiment/.

To provide the robots with the modular controller (M) with a more flexible mechanism for arbitrating between the three modules, we also ran additional experiments in which the time duration of the three phases was encoded in additional evolvable parameters or in which the arbitration between the modules was realised by the robot itself through additional output neurons (as in Nolfi 1997). However, all these experiments led to poorer results with respect to the base (M) condition. The results obtained in these further tests are not included in the paper for reason of space.

The evolutionary algorithm

The initial population consists of 20 randomly generated genotypes, which encode the connection weights and biases of 20 corresponding individual robots (each parameter is encoded by 8 bits and normalised in the range [–5.0, +5.0]). Every generation, each individual is evaluated for three trials in environments that randomly varied in dimension within the limits indicated above. The fitness of each trial is calculated by counting the percentage of 20 × 20 cm portions of the environment that are visited from the robot at least once during the trial. The total fitness is calculated by averaging the fitness obtained during the three trials. All individuals are allowed to generate an offspring that is also evaluated for three trials. The 20 offspring are generated by creating a copy of the parent genotype and by mutating each bit with a 2 % probability. The genotype of offspring is used to replace the genotype of the worst parents or discarded depending on whether or not offspring outperform the parents. The genotypes of the initial population were generated randomly. Each evolutionary experiment was replicated 20 times starting from different randomly generated initial populations.

Results

In “Performance and efficacy of plastic versus non-plastic behavioural solutions”, we describe the performance achieved in the different experimental conditions. As we will see, the cleaning task in the convex environment admits a simple behavioural solution that does not require the exhibition of multiple behaviours. Consequently, the performance obtained in the different experimental conditions is rather similar. On the contrary, the cleaning task in the concave environment requires the exhibition of at least two sub-behaviours that differ in forms and functions: an exploration behaviour that enables the robot to explore the large central area and a wall-following behaviour that enables the robot to explore the peripheral areas and the borders of the central area. The possibility to discover and to display these two behaviours rather than a single undifferentiated behaviour crucially depends on the characteristics of the robots’ neural controller as demonstrated by the fact that the behaviour and the performance significantly vary in the four experimental conditions.

In “On the mechanisms supporting behaviour differentiation and arbitration”, we will discuss the mechanisms that support behavioural differentiation and arbitration by analysing the behavioural solutions found in the different experimental conditions. As we will see, the two most important mechanisms that support the evolution of behavioural plastic solutions are the ability to perceive and to generate affordances (i.e., opportunities for behaviours) and the possibility to flexibly and properly handle behavioural transitions.

Performance and efficacy of plastic versus non-plastic behavioural solutions

By post-evaluating the best robot of the last generation of each replication for 500 trials, we can see how in the concave environment, the evolved robots reach close to optimal performance in the temporal (T) experimental conditions, good performance in the modular conditions (M), and relatively low performance in the case of the simple (S) and range sensor (R) conditions (Fig. 2, left). The performance of each experimental condition statistically differs from all others conditions (Kruskal–Wallis ANOVA, $df = 3$, $p < 0.001$—Bonferroni-corrected Mann–Whitney U, $p < 0.0083$) with the exception of (S) and (R) that do not differ significantly from each other ($p = 0.82$). The performance obtained in the experiments in which the robots were also provided with the internal neurons (Fig. 2, right) does not significantly differ from the experiments without internal neurons (Mann–Whitney U, $p < 0.05$).

The analysis of the behaviours displayed by the best robots of the last generation indicates that the performance level correlates with the ability of the robots to display multiple behaviours. This is clearly illustrated by the behaviour displayed by the best (S) and (T) robots that achieved a fitness of 67.4 and 82.8 %, respectively. While (S) displays a single uniform behaviour along the trial, (T) is capable of performing two well-differentiated behaviours (Fig. 3, top).

Indeed, the best robot with a simple architecture (S) always behaves in the same manner during the successive phases of the trial (Fig. 3, top-left). In particular, it avoids

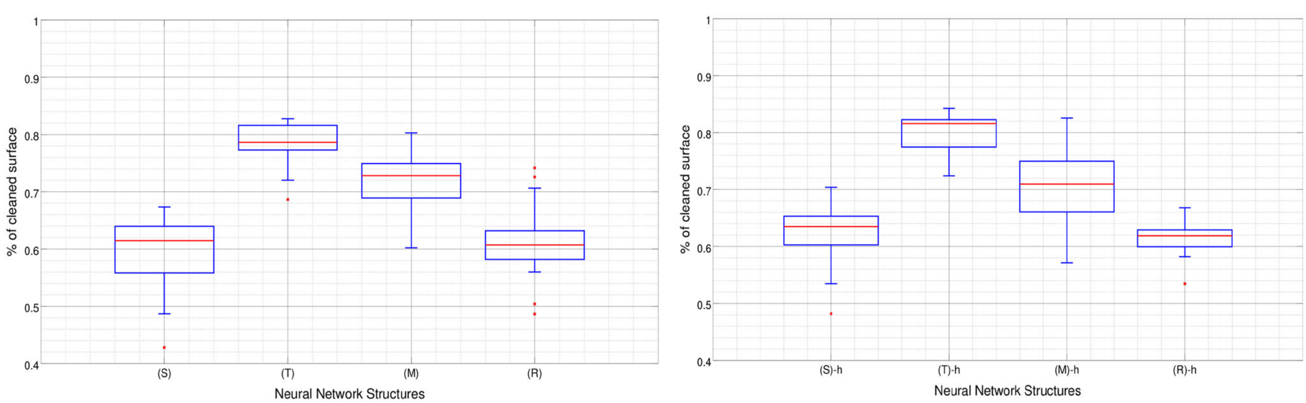

Fig. 2 *Box plots* of performance in the concave environment. The *left* and *right* figures report the results obtained without internal neurons and with internal neurons, respectively. The *box plots* display the performance of the best robot of the last generation in the four experimental conditions, i.e., in the single (S), temporal (T), modular (M), and range sensor (R) conditions. *Boxes* represent the inter-quartile range of the data, while the *horizontal lines* inside the *boxes* mark the median values. The whiskers extend to the most extreme data points within 1.5 times the inter-quartile range from the box. *Circles* mark the outliers. *Each box* displays the performance of the best robot of 20 replications of each experiment. The performance is indicated by the percentage of cleaned cells within the walls. The value corresponding to optimal performance is unknown but is reasonably below 1.0 given that the robots have a rather limited cleaning time

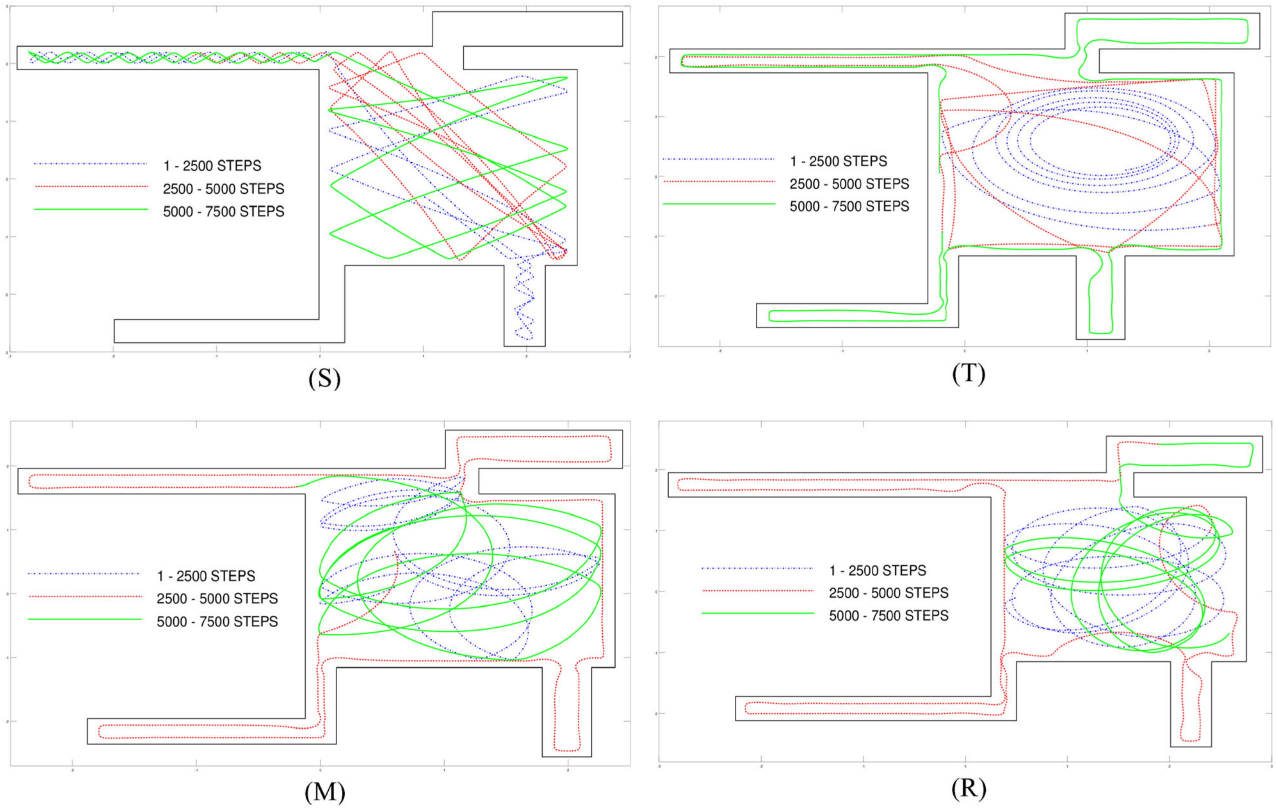

Fig. 3 Typical trajectories displayed by the best robots of the four experimental conditions without hidden units in the concave environment. The portions of the trajectory produced during the first, second, and third part of the trial (i.e., from step 1 to 2500, from step 2501 to 5000, and from step 5001 to 7500, respectively) are shown with different *colours* and *line* style (colour figure online)

walls and obstacles by sharply turning with an angle of 45°–90° (depending on the relative angle with which the robot approaches the obstacle) and moves straight when it is far from obstacles. Through the exhibition of this behaviour, the robot manages to keep exploring the environment until the end of the trial by avoiding obstacles and by keep moving in different portions of the environment. However, the robot spends most of its time by exploring the large central portion of the environment. It explores the peripheral areas only occasionally when it happens to approach them with a direction that it is almost orthogonal to the entrance of the corridor. The robots of the other replications of the experiments show qualitatively similar behaviours (results not shown).

The best robot with the time neuron architecture (T), instead, shows two well-differentiated behaviours: (1) an initial exploration behaviour that is realised by producing a progressively larger curvilinear trajectory that enables the robot to explore the large central portion of the environment, and (2) a wall-following behaviour that enables it to explore all the peripheral areas of the environment (Fig. 3, top-right). Although the way in which the exploration behaviour is realised varies in different replications of the experiment, well-differentiated exploration and wall-following behaviours are clearly observable in all cases (results not shown). The high performance of these robots is due to their ability to display different behaviours, which are specialised for the exploration of large open areas and peripheral areas, and to carefully tune the time duration of the two behaviours. Indeed, the relative duration of the two behaviours determines whether the robot spends enough time exploring the central large area while keeping enough time to explore all the peripheral areas of the environment or not.

A qualitative analysis of the first ten replications showed that in the best two robots, that clearly outperform the best robots of the other eight replications, the transition between the two behaviours occurs at 3.17 $\pm$ 0.11 min. This transition time is optimal or nearly optimal as demonstrated by the fact that post-evaluation tests performed by slowing down or speeding up the robot's internal clock and, consequently, the behaviour transition led to significantly worse performance (results not shown).

The best robot with the modular (M) architecture also shows an exploration behaviour displayed during 4.17 min, when the robot operates on the basis of the first and third neural modules, and a wall-following behaviour displayed during 2.08 min in which it operates on the basis of the second neural module (Fig. 3, bottom, left). The lower performance with respect to the best (T) robot is due to the fact that the transition between the two behaviours is too abrupt and to the fact that it is not able to finely tune the relative duration of the two behaviours. The analysis of the robots of the other replications shows qualitatively similar solutions although, in some cases, the differentiation of the behaviour is less marked (result not shown). As mentioned above, we carried out a series of additional experiments in which the genotype of evolving robots included three additional genes that were used to determine the time duration of the three phases. However, in this condition, the evolved robots relied on a single exploration behaviour, as in the case of the (S) experimental condition (results not shown).

Finally, the analysis of the best robot in the case of the range sensor experimental condition (R) also displays a behavioural plastic solution characterised by the exhibition of an exploration behaviour and a wall-following behaviour (Fig. 3, bottom-right). This robot alternates the two behaviours by switching either from the exploration to the wall-following behaviour or from the wall-following to the exploration behaviour. The achievement of lower performance with respect to the (T) experimental condition is due primarily to the inability of this robot to precisely control the duration of behaviours, as demonstrated by the high variability of the relative duration of the two behaviours among trials. The best robots of four other replications displayed qualitatively similar solution, while the best robot of the five remaining replications display a single uniform exploratory behaviour similar to that shown by (S) robots (result not shown). The behaviour of the second set of ten replications was not inspected.

In the convex environment, instead, the robots achieve similar performance in all experimental conditions (see Fig. 4, left). The differences among the four experimental conditions are significant (Kruskal–Wallis ANOVA, $df = 3$, $p < 0.001$). However, the pairwise comparison (Bonferroni-corrected Mann–whitney U) indicates that this difference is due to the fact that (R) is significantly worse than (T) ($p < 0.001$) and (M) ($p = 0.00143$). All other conditions do not statistically differ ($p > 0.0083$). The performance obtained in the experiments in which the robots were also provided with the internal neurons (Fig. 4, right) does not significantly differ from the basic experiments for (T) and (M) (Mann–Whitney U, $p > 0.05$) with the exception of (S) and (R) in which the performance of the experiments with internal neurons is significantly better in the former, and worse in the latter case (Mann–Whitney U, $p < 0.05$).

Overall, these results can be explained by considering that in this type of environment, the exhibition of a single behaviour is sufficient to achieve close-to-optimal performance. As a consequence, evolving robots do not develop multiple behaviours (see Fig. 5). In some cases, especially in the (M) condition, a weak differentiation is observed. However, it does not provide an advantage in this type of environment.

On the mechanisms supporting behaviour differentiation and arbitration

We have seen how behavioural plasticity, i.e., the ability to display and regulate multiple behaviours, can enable the adaptive robots to achieve better performance in the concave environment and that the emergence of behavioural plastic solutions depends on the characteristics of robot's neural controllers. We will now focus on the mechanisms supporting behaviour differentiation and arbitration. As we will see, evolving robots can rely on different mechanisms to achieve behavioural plasticity. The efficacy of these mechanisms and the facility with which they can be

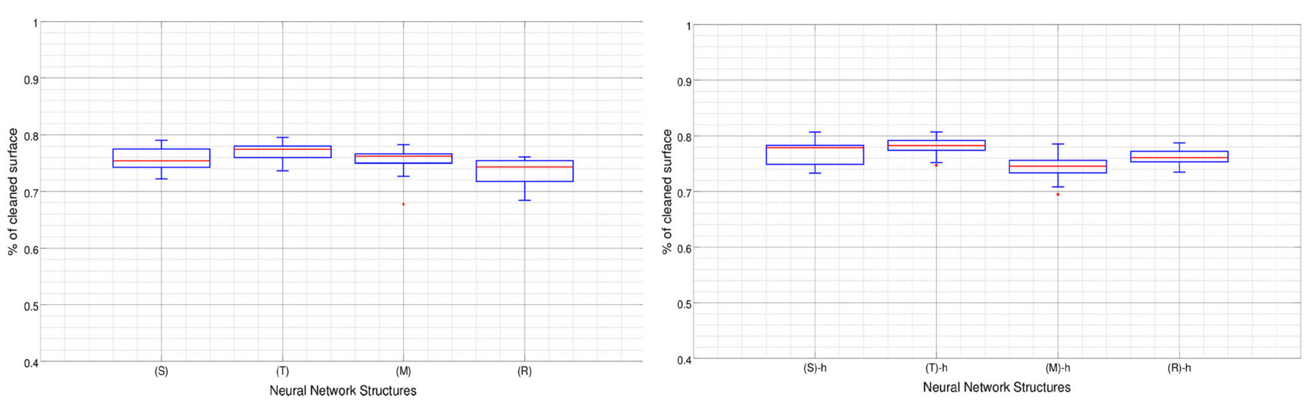

Fig. 4 *Box plots* of performance in the convex environment. The *left* and *right* figures report the results obtained without internal neurons and with internal neurons, respectively

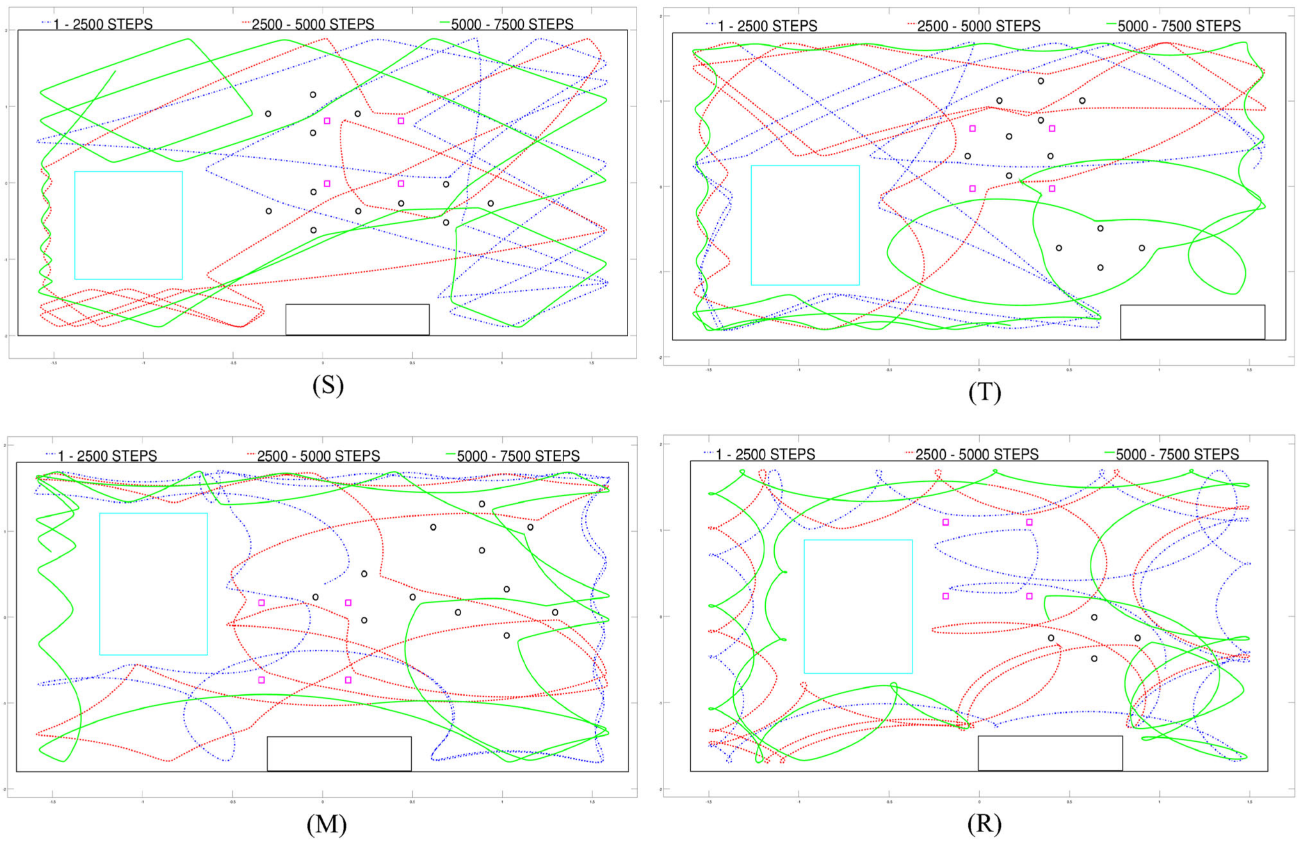

Fig. 5 Typical trajectory displayed by the best robots of the four experimental conditions without hidden units in the convex environment

discovered explain the variations in performance observed in the considered experimental conditions.

Before entering into this, it is important to point out that, as we mentioned in the introduction, the behaviour displayed by an embodied and situated agent is a dynamical process unfolding in time that results from the robot/environmental interactions. This implies that the organisation of behaviour/s varies at different timescales. Moreover, this implies that the sensory states experienced by the robot at a given time step are co-determined by the actions produced by the robot during previous robot/environmental interactions. If we use the term affordance introduced by Gibson (1979) to indicate sensory states that elicit the production of behaviours, this implies that the affordances are not only extracted through sensors from the internal and/or the external environment but are also generated by the robot itself through actions.

The analysis of the behaviour exhibited by the robots at a short timescale (i.e., at a timescale of seconds) indicates that in all experimental conditions, robots tend to exhibit at least two different low-level behaviours: (1) an obstacle-avoidance behaviour that consists in turning while the robot detects an obstacle on its frontal side, and (2) a move-forward behaviour that consists in moving straight or almost straight while the robot does not detect obstacles in its frontal side (see Fig. 6). This implies that at this short timescale, all robots of all experimental conditions display behavioural plastic behaviours. The reasons that explain

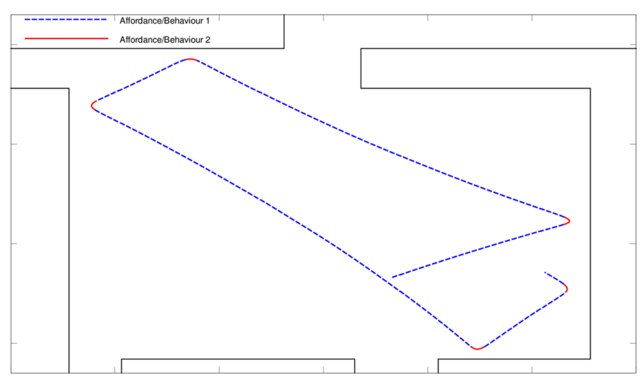

Fig. 6 Exemplification of short-term behavioural plasticity in the case of an exploration behaviour that is realised by alternating a move-forward and an obstacle-avoidance behaviour (shown in *blue* and *red*, respectively). The former behaviour is elicited by perceptual states in which the frontal infrared sensors are not activated, i.e., a state affording the move-forward behaviour. The latter behaviour is elicited by perceptual states in which the frontal infrared sensors are activated, i.e., a state affording the obstacle-avoidance behaviour (colour figure online)

why this type of behavioural plasticity always evolves are that it plays a functional role (i.e., it enables the robot to avoid being stuck and to keep exploring the environment) and that it is supported by the availability of always-available and easy-to-use affordances. Indeed, independently from the way in which the robot behaves, it will always experience a lack of activation on the frontal infrared sensors when the robot/environment context affords a move-forward behaviour and an activation on the frontal infrared sensors when the robot/environmental context affords an obstacle-avoidance behaviour. The infrared sensors, therefore, always enable the robot to perceive when the former or the latter behaviour should be produced and when the transition between the two behaviours should occur.

This ideal situation, however, in which the robot can rely on robust and ready-to-use affordance states only characterises few lucky cases (incidentally, this probably explains why the combination of obstacle-avoidance and navigation behaviours represents a widely used experimental scenario in robotics). In other cases, the affordance states supporting behaviour differentiation and arbitration should be extracted through internal elaboration and/or generated through the exhibition of appropriate behaviours.

This also implies that plasticity is not a binary but rather a continuous property. The greater the number of behaviours/complexity of the sub-behaviours exhibited by a robot is and the greater is the range of timescales at which the robot exhibits differentiated behaviours, the greater the behavioural plasticity of the robot is. In the rest of the paper, however, we focus exclusively on the longer timescale. Consequently, we use the term multiple behaviours and behavioural plasticity to indicate robots that exhibit behaviour differentiation at this timescale, independently of whether they show behaviour differentiation at shorter timescale. We do this since at longer timescale, we observe qualitatively and quantitatively different solutions in the context of our experiments.

As we have seen in the previous section, the concave environment requires behavioural diversification at the longer timescale, e.g., it requires the exhibition of an exploration and a wall-following behaviour lasting for minutes. In this case, however, the robot cannot rely on ready-to-use affordances that indicate when the robot should display the first or the second behaviour and when the robot should switch from one to the other behaviour. To achieve this kind of behavioural plasticity, the evolving robots should find a way to: (1) keep producing the same behaviour for a prolonged period of time, (2) switch behaviour at the right moment, and (3) realise a suitable transition during behaviour switch. We will illustrate in details how the evolved robots manage to master these requirements in the different experimental conditions in the next three sub-sections.

Notice that the evolution of context-dependent behaviours requires the concurrent development of two interdependent skills, the ability to produce a new behaviour and the ability to regulate appropriately when the new behaviour should be exhibited (Williams 1966; West-Eberhard 2003). We will come back on this issue in the concluding section.

Producing behaviours for prolonged periods of time

All evolved robots solve the problem of producing a given behaviour for a prolonged period of time by realising each behaviour in a way that ensures that they keep experiencing stimuli of the right type during the execution of that behaviour. In cases in which the robots should exhibit two differentiated behaviours, i.e., an exploration and a wall-following behaviour, this implies that they should realise the former and the latter behaviours in a way that ensures that they keep experiencing stimuli of type 1 and 2 while they exhibit the former or the latter behaviour, respectively, and should react to the stimuli of the two types by producing actions that enable them to keep producing the former or the latter behaviours, respectively. The two classes of stimuli, thus, assume the role of affordance for the first and for the second behaviours, respectively. These affordances are not directly available from the environment, as in the case of the states affording the obstacle-avoidance and move-forward behaviour discussed above, but are generated by the robots themselves through their actions (i.e., through the ability to realise each behaviour in a way that ensures that the robot keeps experiencing the corresponding affordances). This form of dynamical stability presents some similarities with the one that can be obtained in situated agents through homeokinesis (Der and Martius 2012), a task-independent learning process that can enable situated robot to synthesise temporarily stable behaviours, though the mechanism and the processes through which this is realised are completely different.

All robots displaying multiple behaviours (i.e., (R), (M) and (T) robots) exploit this affordance generation mechanism. However, the (T) and some of the (M) robots also exploit other additional mechanisms that enable the robots to keep producing each behaviour for a prolonged period of time. Thus, let us start by describing the strategy used by the best (R) robot that only relies on this affordance generation mechanism.

The best (R) robot realises the exploration behaviour by moving forward far from obstacles and by turning left near obstacles located in its frontal and frontal-right side and realises the wall-following behaviour by moving forward along walls when it perceives an obstacle on its left side and by turning left when the activations of its left-side sensors decrease (see Fig. 3, bottom left). By behaving in this way, the robot ensures that it keeps experiencing sensory states of type 1 during the exploration behaviour and sensory states of type 2 during the wall-following behaviour (where type 1 includes states in which the infrared sensors are not activated or in which the frontal or right infrared sensors are activated and type 2 includes states in which the left infrared sensors are activated). In other words, as we said above, the problem of keep producing the two behaviours for prolonged period of time is solved by producing each behaviour in a way that ensures that the robot keeps experiencing stimuli affording the same behaviour (i.e., stimuli that elicit actions which lead to the production of the same behaviour).

In (M) robots, the problem of producing the same behaviour for a prolonged period of time is solved also through the development of neural modules specialised for the production of the exploration or of the wall-following behaviour. However, (M) robots rely on affordance generation as well. Indeed, even in some (M) robots, the same neural module enables the robot to produce either the exploration or the wall-following behaviour and to keep producing the current behaviour for a prolonged period of time (Fig. 8). In these cases, the behaviour that is initially triggered depends on the initial position of the robot (i.e., depends on the behaviour afforded by the first experienced sensory states).

In (T) robots, the cue provided by the temporal neuron co-determines the behaviour produced by the robot and, consequently, is used to keep producing the current behaviour for a prolonged period of time. Indeed, whether the robot keeps producing the exploration behaviour or switches to the wall-following behaviour also depends on the state of the temporal neuron (see Fig. 9). On the other hand, the state of the time neuron influences the duration of the exploration behaviour only during a critical phase, i.e., when the state of the time neuron is smaller than 0.6 and greater than 0.4. During the rest of the trial, the ability of the robot to keep producing the exploration behaviour or the wall-following behaviour relies on an affordance generation mechanism analogous to that described above for the best (R) robot. Interestingly, in the case of the best (T) robot, the temporal neuron is also used to progressively vary over time the way in which the exploration behaviour is realised so as to regulate the probability that the robot keeps experiencing sensory state affording the execution of this behaviour. Indeed, by initially moving forward and turning left of several degrees, the robot eliminates, completely, the possibility to encounter a wall on its left side (i.e., the possibility to experience stimuli affording the alternative wall-following behaviour). Then, by moving forward and progressively reducing the angle of turn over time, the robot becomes progressively kinder with respect to the possibility of experiencing stimuli affording the wall-following behaviour. This brings us to the question of how robots manage to switch behaviour.

Switching between alternative behaviours

The problem of switching between different behaviours is also solved through affordance generation. To understand how robots can act in a way that enables them to both experience stimuli affording the current behaviour and stimuli affording the alternative behaviour, we should reformulate the definition of affordance generation in probabilistic terms. Evolved robots solve the problem of producing a given behaviour for a prolonged period of time and the problem of switching behaviour by realising behaviours in a way that ensures that they keep experiencing stimuli affording the current behaviour with a given high probability and stimuli affording the alternative behaviour with a given low probability, respectively.

All evolved robots solve the problem of keep producing the same behaviour for a prolonged period of time and the problem of switching behaviour in this way. However, some robots also rely on additional complementary mechanisms, as we illustrate below.

In the case of the best (R) robot, the switches from the exploration behaviour to the wall-following behaviour occur when the robot encounters a wall on its frontal-left side during the execution of the exploration behaviour (see Fig. 8, left), a situation that occurs with a low probability for the reason described in the previous section. Overall, this means that the exploration behaviour is realised in a way that the robot keeps experiencing stimuli affording the exploration behaviour most of the time, while occasionally experiencing stimuli affording the alternative behaviour. Clearly, this is an example of how the simultaneous evolution of form and regulation can be solved. The same mechanism is responsible for behaviour production (i.e., the prolonged production of the same behaviour) and for behaviour switch. The fact that this solution is never found by (S) robots indicates that the availability of the additional cues provided by the range sensors enables (R) robots to regulate, more effectively, the probability with which the robots experience stimuli affording the current or the alternative behaviour. This affordance generation strategy enables the best (R) robot to switch from the exploration to the wall-following behaviour at the optimal moment on the average but with a high variability among trials (the robot switches at 2.99 ± 1.02 min). The high variability negatively impacts on performance since it often leads to situations in which the time dedicated to the two behaviours is unbalanced. The problem is particularly serious when the switch from the exploration behaviour to the wall-following behaviour occurs too early, since circling along the periphery of the environment for more than one lap is useless. This probably explains why the best robot of the (R) experimental condition also developed an ability to switch back from the wall-following behaviour to the exploration behaviour when the robot encounters a wall frontally after exiting from a peripheral corridor (see Fig. 7, right). This latter ability is lacking in the best robots of the other replications that consequently achieve lower performance. In other words, the best (R) robot is capable of displaying reversible behavioural switch.

In the case of the robot evolved in the (M) experimental conditions, in which the three neural modules control the robot during three successive phases of 2.08 min, not surprisingly behavioural switching occurs primarily during the switch between the first and the second module and/or between the second and the third module. The rigidity of this mechanism, however, does not enable the robot to regulate the exact moment in which the switch is realised. In most of the replications, the exploration behaviour is produced for 4.17 min and the wall-following behaviour for 2.08 min since two modules specialise for the production of the former behaviour and the remaining module specialises for the production of the latter behaviour. However, also these robots use affordance generation to switch between behaviours. Indeed, as we mentioned in the previous section, some of the best (M) robots also display an ability to switch behaviour while they operate on the basis of the same neural module through the same affordance generation mechanism described above (see Fig. 8). The usage of this strategy enables these robots to achieve a more balanced allocation of time to the two behaviours that, in turn, enables it to achieve better performance with respect to the best robots of the other replications.

In the case of the robot evolved in the (T) experimental condition, the switch is regulated by both the stimuli experienced by the robot (i.e., by affordance generation) and by the cue provided by the robot's internal clock. This double regulation enables the best (T) robot to carefully balance the time allocated to the two types of behaviour and to reduce the variability among trials (i.e., the transition occurs 3.17 ± 0.11 min). The double regulation

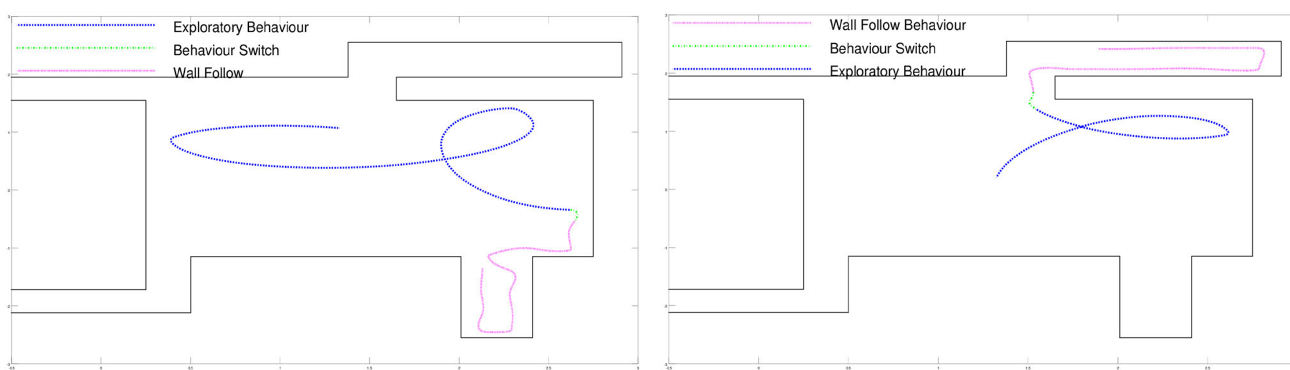

Fig. 7 Illustration of how the best (R) robot switches from the exploration to the wall-following behaviour and vice versa (*left* and *right*, respectively)

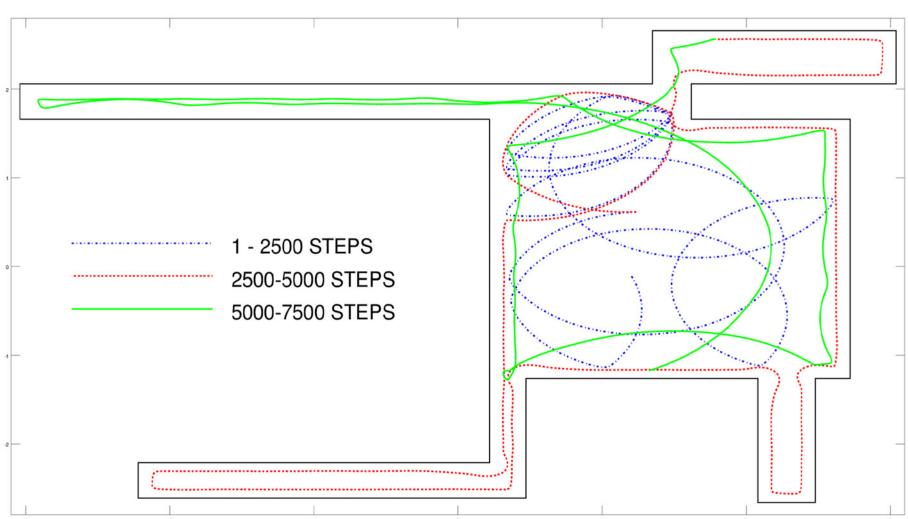

Fig. 8 Trajectory produced by the best (M) robot which produces an exploration behaviour under the control of the first neural module, an exploration and then a wall-following behaviour under the control of the second neural module, and a wall-following and then an exploration behaviour under the control of the third neural module

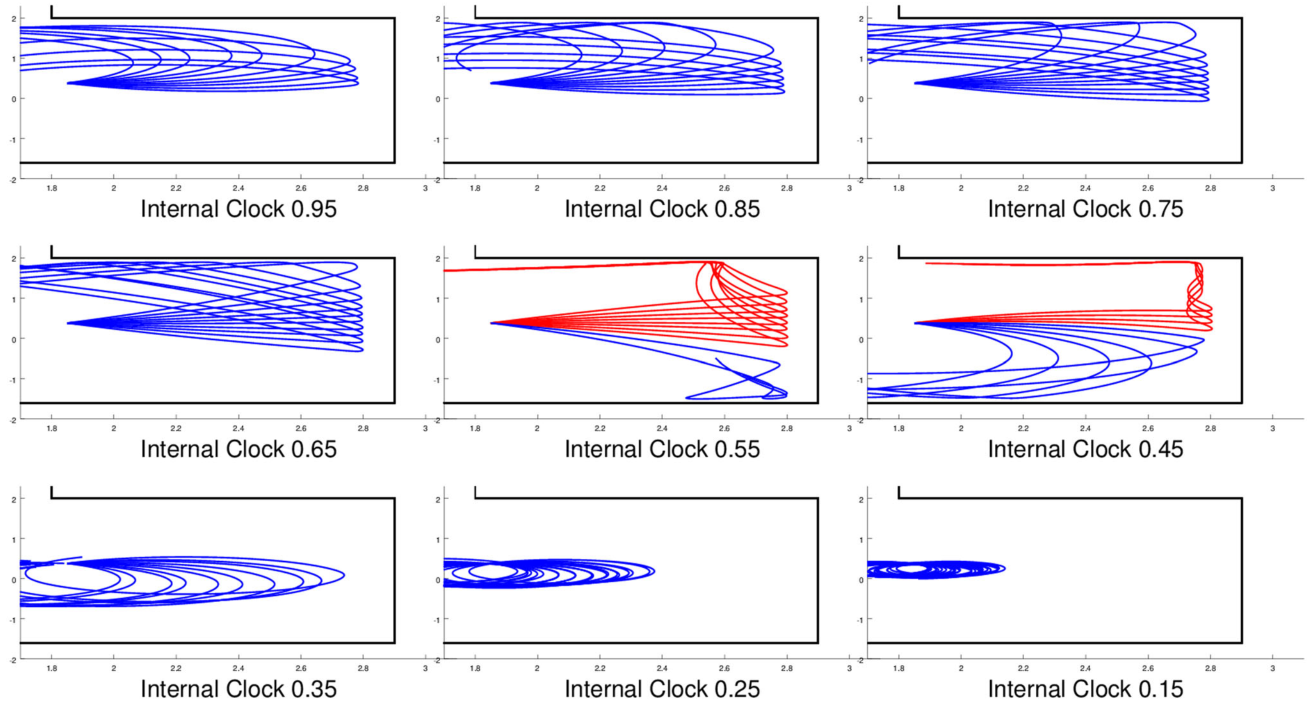

Fig. 9 Behaviour produced by the best (T) robot during different trials in which it started from the same initial position with systematically varied orientations and systematically varied state of the time neuron. The *red* and *blue lines* represent the trajectories produced by the robot during trials in which it switches or does not switch to the wall-following behaviour, respectively. The *black lines* represent the walls. For sake of clarity, we only show the local portion of the environment in which the robot is located (colour figure online)

process is demonstrated by the analysis of the trajectories produced by the robot during a series of trials in which the robot always starts from the same position and in which the orientation of the robot and the state of the time neuron are systematically varied (Fig. 9). As shown in Fig. 9, whether the robot switches or not to the wall-following behaviour depends both on the state of the internal clock and on the state of the infrared sensor that the robot experiences when it approaches the wall. Overall, this shows that whether the switch between the two behaviours occurs or not depends both on the state of the internal clock and on the way in which the exploration behaviour is realised which, in turn, influences the type of stimuli that the robot experiences. As mentioned above, in the case of the best (T) robot, the state of the time neuron is not only used to regulate the probability that the robot switches behaviour directly (the probability that the robot initiates a wall-following behaviour in a given relative position in the environment) but is also used to regulate the way in which the exploration behaviour is realised which, in turn, influences the probability that the robot will later experience stimuli affording the wall-following behaviour.

Realising suitable and effective behaviour transitions

The connectedness of behaviours, i.e., the fact that alternative behaviours are semi-discrete and semi-dissociable units that are only partially independent, implies that the transitions between behaviours should be handled with care. In the case of our experiments, in particular, the transition between the exploration and the wall-following behaviour requires special care since the latter behaviour can only be produced when the robot is located near a wall and when the wall to be followed is located on a specific side of the robot. Indeed, the analysis of the evolved robots shows that the way in which the behaviour transitions are handled in evolved robots has an important impact on robots' performance.

The transition problem is particularly severe in the (M) experimental condition when the behavioural switch typically occurs suddenly after 2.08 and 4.16 min as a result of the neural module switch. The problem is so severe that in three out of the first ten replications, the second control module specialises simply for handling the transition (Fig. 10). In other words, these robots dedicate the second 2.08-min phase simply to move towards a location from which the wall-following behaviour can be effectively initiated.

The smartest solution to the transition problem is that discovered by the best (T) robot (see Fig. 3, right). Indeed, as we mentioned above, this robot exploits the cue provided by the internal clock to gradually modify the exploration behaviour so as to ensure that the robot will always reach a relative location with respect to the walls

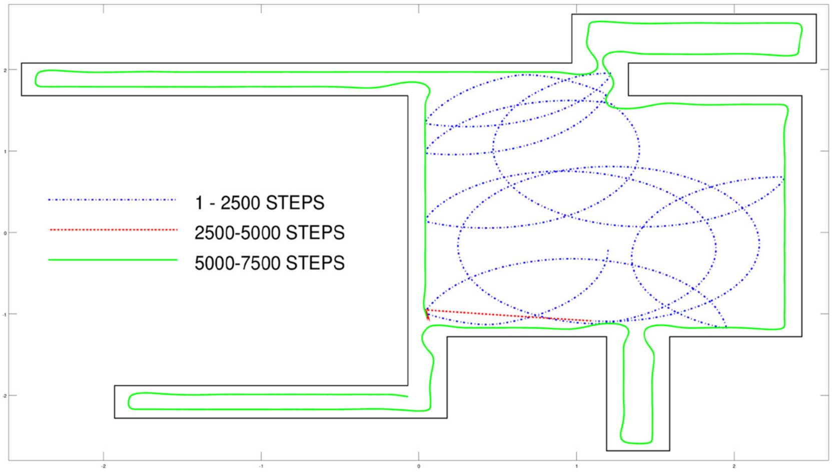

Fig. 10 Trajectory produced by one of the three best (M) robots characterised by a second module that is specialised for enabling a suitable transition from the exploration to the wall-following behaviour

from which the wall-following behaviour can be effectively triggered during the critical period (i.e., during 3.17 $\pm$ 0.11 min). Overall, this leads to an extremely timely, smooth and effective transition that enables this robot to outperform all other robots.

Conclusions

In this paper, we demonstrated that behavioural plasticity can evolve in artificial robots, independently from whether the task does or does not require to face multiple conflicting goals. Indeed, the solution of a task involving a single objective (e.g., cleaning a given area) can also benefit from the utilisation and the combination of multiple differentiated behaviours. Moreover, we demonstrated how the exploitation of behavioural plasticity enables evolving robots to achieve better performance.

Interestingly, the behaviours displayed by the best evolving robots show similarities with those obtained by Gordon et al. (2014) through a minimal model based on intrinsic motivation in which novelty is used as an intrinsic reward. More generally, the adaptive advantage provided by the ability to display multiple behaviours suggests that a potential benefit of task-independent fitness functions, which encourage the development of novel behaviours (see Schmidhuber 1990; Oudeyer et al. 2007; Martius et al. 2014), consists in facilitating the synthesis of behavioural plastic solutions.

The analysis of the obtained results indicates that the mechanisms that support the evolution of behavioural plastic solutions are the ability to perceive affordances (i.e., perceptual states encoding opportunities for behaviours) and the ability to realise smooth and effective transitions between different behaviours.

The perception of affordance constitutes a prerequisite for the possibility to develop differentiated behaviour and for the possibility to effectively arbitrate them, i.e., selecting the behaviour that is appropriate for the current robot/environmental context and regulating the duration of each behaviour. Interestingly, the basic mechanism that is used by evolving robots to perceive affordances is affordance generation, i.e., the ability to realise each behaviour in a way that ensures that the robot keeps experiencing sensory state affording the current behaviour with a given high probability and sensory states affording alternative behaviours with a given low probability.

The limitations of this affordance generation mechanism, e.g., the inability to finely tune the duration of behaviours, are overcome by using additional regulatory processes that rely on internal cues. In particular, in the case of the best evolved robot, this is realised by complementing the basic affordance generation mechanism with two additional regulatory processes. The second additional regulatory process consists in using the state of the internal clock to progressively vary the way in which the exploration behaviour is realised so as to progressively increase the probability that the robot will experience stimuli affording the wall-following behaviour (see Fig. 3, top-right). The third additional regulatory process consists in using the state of the internal clock to vary qualitatively the way in which the robot reacts to perceived stimuli (e.g., to avoid obstacles by turning right or left which causes the robot to later perceive stimuli affording the exploration behaviour or the wall-following behaviour, respectively, see Fig. 9).

Overall, this implies that behaviour arbitration in the best evolved robots is realised through the combined effects of multiple partially redundant regulatory processes that operate through weak interactions. This type of organisation is advantageous both from an evolutionary perspective, since it enables a gradual transformation (Conrad 1990; Krischner and Gerhart 2005), and from a functional perspective, since it enables the robots to operate on the basis of the combined effect of multiple factors. This type of organisation might, indeed, be crucial to enable the concurrent evolution of form and regulation (sub-behaviours and behaviour arbitration in the case of our experiment).

While the importance of affordance perception and usage is widely recognised, the notion of affordance generation that we introduced in this paper and the description of how affordance generation supports the evolution of multiple context-dependent behaviour are original, to our knowledge.

The need to realise smooth and effective transitions between behaviours originates from the fact that behaviour is a dynamical process in which the state of the system at time t critically influences the state of the system at time $t + 1$. In other words, it originates from the fact that the way in which a first behaviour is realised influences the way in which the second following behaviour is realised. More generally, this implies that, as claimed by West-Eberhard (2003), the modular organisation of behaviour is characterised by subunits that are semi-discrete and semi-dissociable, i.e., that are not fully separable and dissociable.

Also, from this perspective, the possibility to operate on the basis of multiple regulatory processes, such as those described above, presents important advantages. In particular, the affordance generation mechanism that exploits the sensory state currently experienced by the robot to determine the behaviour to be exhibited ensures that the behaviour exhibited by the robot is always appropriate to the current robot/environmental context. On the other hand, the regulation processes, carried out on the basis of the state of the robot's internal clock, ensure that behavioural switch will occur within the appropriate time window.

In robotics, the objective of designing robots capable of displaying elaborate behaviours is usually pursued by designing modular controllers, eventually organised hierarchically, in which each module is specialised for the production of a corresponding sub-behaviour, and in which modules are alternated on the basis of some arbitration mechanism (Brooks 1986; Stone and Veloso 2000; Van Hoorn et al. 2009). In these works, the decomposition of the overall behaviour into sub-behaviours and, consequently, the organisation of the modules, are usually designed by the experimenter, while in other cases, it is learned (Tani and Nolfi 1999; Haruno et al. 2001). From this perspective, our results suggest that the utilisation of behaviour generation and arbitration mechanisms that are rigid and/or that do not support the realisation of smooth and effective behaviour transitions might constitute a strong limitation.

Acknowledgments This work is partially funded by CAPES through the Brazilian program Science Without Borders. The authors thank Vito Trianni, Luca Simione, and the anonymous reviewers for useful comments on the first drafts of the paper.

References

Ackerman E (2010) Long exposure pictures of robots cleaning [Blog post]. http://spectrum.ieee.org/automaton/robotics/home-robots/long-exposure-pictures-of-robots-cleaning. Retrieved April 23, 2015

Bangard J (2011) Spontaneous evolution of structural modularity in robot neural network controller (2011). In: Proceedings of the 13th annual conference companion on genetic and evolutionary computation (GECCO), ACM, New York

Barlow GW (1977) Modal action patterns. In: Sebeok TA (ed) How animals communicate. Indiana University Press, Bloomington

Bonani M, Longchamp V, Magnenat S, Retornaz P, Burnier D, Roulet G, Vaussard F, Bleuler H, Mondada F (2010) The marXbot, a miniature mobile robot opening new perspectives for the collective robotic research. In: 2010 IEEE/RSJ international conference on intelligent robots and systems, pp 4187–4193

Brooks RA (1986) A robust layered control system for a mobile robot. Robot Autom 2(1):14–22

Calabretta R, Nolfi S, Parisi D, Wagner G (2000) Duplication of modules facilitates the evolution of functionalspecialization. Artif Life 6(1):69–84

Chemero A (2011) Radical embodied cognitive science. MIT Press, Cambridge

Conrad M (1990) The geometry of evolution. BioSystems 24:61–81

Der R, Martius G (2012) The playful machine: theoretical foundation and practical realization of self-organizing robots. Springer, Berlin

Eberhard WG (1988) Behavioural flexibility in orb web construction: effect of supplies in different silk glands and skider size and weight. J Arachnol 16:295–302

Fentress JC (1983) A view of ontogeny. In: Esenberg J, Kleiman D (eds) Special publications American Society of Mammalogists, vol 7, pp 24–64

Floreano D, Nolfi S (1997) Adaptive behavior in competing co-evolving species. In: Husband P, Harvey I (eds) Proceedings of the fourth conference on artificial life. MIT Press, Cambridge

Gallistel CR (1980) The organization of action. A new synthesis. Lawrence Erlbaum, Hillsdale

Gibson J (1979) The ecological approach to visual perception. Houghton-Mifflin, Boston

Gordon G, Fonio E, Ahissar E (2014) Emergent exploration via novelty management. J Neurosci 34(38):12646–12661

Haruno M, Wolpert DM, Kawato M (2001) Mosaic model for sensorimotor learning and control. Neural Comput 13:2201–2220

Hinde RA (1970) Animal behavior. McGraw Hill, New York

Huizinga J, Mouret B, Clune J (2014) Evolving neural networks that are both modular and regular: hyperneatplus the connection cost technique. In: Proceedings of the genetic and evolutionary computation conference (GECCO), ACM, New York

IRobot (2013) Our history. http://www.irobot.com/about-irobot/company-information/history.aspx. Retrieved April 29, 2015

Izquierdo E, Bührmann T (2008) Analysis of a dynamical recurrent neural network evolved for two qualitatively different tasks: walking and chemotaxis. In: Bullock S, Noble J, Watson RA, Bedau MA (eds) Proceedings of the 11th international conference on the synthesis and simulation of living systems (ALIFE 11). MIT Press, Cambridge

Jackson RR, Wilcox RS (1993) Spider flexibly chooses aggressive mimicry signals for different prey by trial and errors. Behavior 127:21–36

Komers PE (1997) Behavioural plasticity in variable environments. Can J Zool 75:161–169

Krischner MW, Gerhart JC (2005) The plausibility of life: resolving Darwin's Dilemma. Yale University Press, USA

Martius G, Der R, Herrmann JM (2014) Robot learning by guided self-organization. Springer, Berlin

Massera G, Ferrauto T, Gigliotta O, Nolfi S (2013) FARSA: an open software tool for embodied cognitive science. In: Lio' P, Miglino O, Nicosia G, Nolfi S, Pavone M (eds) Proceeding of the 12th European conference on artificial life. MIT Press, Cambridge

Mitchell SD (1990) The units of behaviour in evolutionary explanations. In: Bekoff M, Jamieson D (eds) Interpretation and Explanation in the Study of Animal Behavior. Westview Press, Boulder

Montes Gonzalez F, Prescott TJ, Gurney K Humphries M, Redgrave P (2000) An embodied model of action selection mechanisms in the vertebrate brain. In: Meyer J-A, Berthoz A, Floreano D, Roitblat H, Wilson SW (eds) From animals to animats 6: Proceedings of the sixth international conference on simulation of adaptive behaviour. MIT Press, Cambridge

Nolfi S (2009) Behavior and cognition as a complex adaptive system: insights from robotic experiments. In: Hooker C (ed) Handbook of the philosophy of science: philosophy of complex systems, vol 10. General editors: Dov M. Gabbay, Paul Thagard, John Woods. Elsevier

Nolfi S, Floreano D (1999) Learning and evolution. Auton Robot 7:89–113

Nolfi S, Floreano D (2000) Evolutionary robotics: the biology, intelligence, and technology of self-organizing machines. MIT Press/Bradford Books, Cambridge

Nolfi S, Parisi D (1997) Learning to adapt to changing environments in evolving neural networks. Adapt Behav 1:75–98

Nolfi S, Bongard J, Husband P, Floreano D (2016) Evolutionary robotics. In: Siciliano B, Khatib O (eds) Handbook of robotics, Springer, Berlin **(in press)**

Otte T (1972) Simple versus elaborate behavior in grasshoppers. An analysis of communication in the genus *Syrbula*. Behaviour 42:291–322

Oudeyer P-Y, Kaplan F, Hafner V (2007) Intrinsic motivation systems for autonomous mental development. IEEE Trans Evol Comput 11(2):265–286

Petrosino G, Parisi D, Nolfi S (2013) Selective attention enables action selection: evidence from evolutionary robotics experiments. Adapt Behav 21(5):356–370

Prescott TJ (2008) Action selection. Scholarpedia 3(2):2705

Rahim SA, Yusof AM, Braunl T (2014) Genetically evolving action selection mechanisms in a behavior-based system for target tracking. Neurocomputing 133:84–94

Schmidhuber J (1990) A possibility for implementing curiosity and boredom in model-building neural controllers. In: From animals to animals: proceedings of the first international conference on simulation of adaptive behavior. MIT Press, Cambridge

Schrumand J, Miikkulainen R (2012) Evolving multimodal networks for multitask games. IEEE Trans Comput Intell AI Games 4(2):94–111

Seth A (2012) Optimized agent based modelling of action selection. In: Seth A, Prescott TJ, Brysonn JJ (eds) Modelling natural action selection. Cambridge University Press, Cambridge

Seth A, Prescott TJ, Brysonn JJ (2012) Modelling natural action selection. Cambridge University Press, Cambridge

Stone P, Veloso M (2000) Layered learning. In ECML. Springer, Berlin, pp 369–381

Tani J, Ito M (2007) Self-organization of behavioural primitives as multiple attractor dynamics: a robot experiment. IEEE Trans Syst Man Cybern. Part A Syst Hum 33(4):481–488

Tani J, Nolfi S (1999) Learning to perceive the world as articulated: an approach for hierarchical learning in sensory- motor systems. Neural Netw 12:1131–1141

Van Hoorn N, Togelius J, Schmidhuber J (2009) Hierarchical controller learning in a first-person shooter. In CIG. IEEE, pp 294–301

Verbancsics P, Stanley KO (2011) Constraining connectivity to encourage modularity in HyperNEAT. In: Proceedings of the 13th annual conference companion on genetic and evolutionary computation (GECCO), ACM, New York

Wenzel JW (1993) Application of the biogenetic law to behavioural ontogeny: a test using nest architecture in paper wasps. J Evol Biol 6:229–247

Williams GC (1966) Adaptation and natural selection. Princeton University Press, Princeton

West-Eberhard MJ (2003) Developmental plasticity and evolution. Oxford University Press, New York

Williams P, Beer R (2013) Environmental feedback drives multiple behaviors from the same neural circuit. In Advances in artificial life, ECAL 2013, vol 12, pp 268–275