2017

Adaptive Behaviour in Evolving Robots

Tyska Carvalho, Jonata

http://hdl.handle.net/10026.1/10547

http://dx.doi.org/10.24382/629
University of Plymouth

Adaptive Behaviour in Evolving Robots

by

Jônata Tyska Carvalho

A thesis submitted to the Plymouth University
in partial fulfillment for the degree of

Doctor of Philosophy

School of Computing, Electronics and Mathematics
ISTC-CNR PhD Node

December 2017

Acknowledgements

It is surreal for me to be writing this. What seemed so distant some years ago, now is so close to the end. I have so many people to be grateful to that I don't even know where to start. I really own a lot to these people that helped to forge the person who I am, personally and professionally, from the beginning of my life to the present day.

First of all I would like to thank my supervisor Stefano Nolfi for all the knowledge, support and patience he provided me along these years. I thank also all my colleagues from ISTC/LARAL (too many, I'm not putting names here) that certainly helped me technically, but much more important than that was the friendship and the amazing environment they helped to create in the last four years. I am also grateful to my second supervisor Angelo Cangelosi and all the Plymouth University staff for the amazing support they provided even being far away.

I am thankful to my work colleagues from Centro de Ciencias Computacionais(C3)/FURG in Brazil, that trusted me and allowed me to stay these 4 years away for finishing my PhD. Especially those who visited me in this period, you can't imagine how important this was for me to keep going.

I would like to thank my Brazilian best friends Marcel and Cristofer, and my brother Jaderson, which even being far away had the important mission to keep me connected with the part of myself that remained in Brazil. Impossible not to thank also my Italian best friend Simone, which, even being a bad person, made my stay in Italy much better.

Special thanks to my wife Lara, one of the most amazing person that I've ever met in my life, for all the support and love she provided me so far. I love her with all my heart, and I'm sure she knows that. And finally, I am grateful to my parents Cyryla and Paulo, that gave me everything I needed in this life, even when that meant abandoning some of their dreams. They are my example, and I will always love them for that.

Abstract

In this thesis, the evolution of adaptive behaviour in artificial agents is studied. More specifically, two types of adaptive behaviours are studied: articulated and cognitive ones. Chapter 1 presents a general introduction together with a brief presentation of the research area of this thesis, its main goals and a brief overview of the experimental studies done, the results and conclusions obtained. On chapter 2, I briefly present some promising methods that automatically generate robot controllers and/or body plans and potentially could help in the development of adaptive robots. Among these methods I present in details evolutionary robotics, a method inspired on natural evolution, and the biological background regarding adaptive behaviours in biological organisms, which provided inspiration for the studies presented in this thesis. On chapter 3, I present a detailed study regarding the evolution of articulated behaviours, i.e., behaviours that are organized in functional sub-parts, and that are combined and used in a sequential and context-dependent way, regardless if there is a structural division in the robot controller or not. The experiments performed with a single goal task, a cleaning task, showed that it is possible to evolve articulated behaviours even in this condition and without structural division of the robot controller. Also the analysis of the results showed that this type of integrated modular behaviours brought performance advantages compared to structural divided controllers. Analysis of robots' behaviours helped to clarify that the evolution of this type of behaviour depended on the characteristics of the neural network controllers and the robot's sensorimotor capacities, that in turn defined the capacity of the robot to generate opportunity for actions, which in psychological literature is often called affordances. In chapter 4, a study seeking to understand the role of reactive strategies in the evolution of cognitive solutions, i.e. those capable of integrating information over time encoding it on internal states that will regulate the robot's behaviour in the future, is presented. More specifically I tried to

understand whether the existence of sub-optimal reactive strategies prevent the development of cognitive solutions, or they can promote the evolution of solutions capable of combining reactive strategies and the use of internal information for solving a response delayed task, the double t-maze. The results obtained showed that reactive strategies capable of offloading cognitive work to the agent/environmental relation can promote, rather than prevent the evolution of solutions relying on internal information. The analysis of these results clarified how these two mechanisms interact producing a hybrid superior and robust solution for the delayed response task.

Author's declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award. Work submitted for this research degree at Plymouth University has not formed part of any other degree either at Plymouth University or at another establishment.

Relevant scientific seminars and conferences were regularly attended at which this work was presented. Two articles have been accepted for publication in refereed journals and two papers in international conferences.

This study was carried out in collaboration with the Istituto di Scienze e Tecnologie della Cognizione (ISTC) - Consiglio Nazionale delle Ricerche (CNR), Rome. It was financed by the Brazilian Program Science without Borders through CAPES(Ministry of Education).

This thesis contains works which were the result of collaborations with other researcher. The author contribution to the reported works over the total was about 80% for the work described in chapter 3 and 80% for the work described in chapter 4.

Word count for the main body of this thesis: **33032**

Publications

Carvalho, J.T. & Nolfi, S. *Behavioural Plasticity in Evolving Robots.* Theory in Biosciences 2016, Volume 135, Issue 4, pp. 201-216

Carvalho, J.T. & Nolfi, S. *Cognitive Offloading does not Prevent but Rather Promotes Cognitive Development.* PLOS One 2016, 11:8

Carvalho, J.T. & Nolfi, S. *Functional Modularity Enables the Realization of Smooth and Effective Behavior Integration.* Proceedings of the Artificial Life Conference 2016, Cancun, 2016.

Carvalho, J.T. & Nolfi, S. *Affordance Generation Enables Behavioral Plasticity and Cognitive Offloading in Evolving Robots.* 2016 IEEE Symposium Series on Computational Intelligence (SSCI), Athens, 2016.

Carvalho, J.T. & Nolfi, S. *Exploiting Environmental Differentiation to Promote Evolvability in Artificial Evolution.* GECCO '17 Companion- Proceedings of the 2017 on Genetic and Evolutionary Computation Conference Companion, Berlin, 2017.

Milano, N., Carvalho, J.T. & Nolfi, S. *Environmental Variations Promote Adaptation in Artificial Evolution.* 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, 2017.

Carvalho, J.T. & Nolfi, S. *Favoring the Evolution of Adaptive Robots Through Environmental Differentiation.* 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, 2017.

Attended Conferences, Workshops and Summer schools

AISC: Bodies, Tools & Cognition Conference AISC: Bodies, Tools & Cognition Conference 2014, Rome. Participant, 2014

ISAL: Artificial Life Proceedings of the Artificial Life Conference 2016, Cancun-Mexico. Oral presentation: *Functional Modularity Enables the Realization of Smooth and Effective Behavior Integration.* 2016.

IEEE: Symposium Series on Computational Intelligence 2016 IEEE Symposium Series on Computational Intelligence (SSCI), Athens. Oral presentation: *Affordance Generation Enables Behavioral Plasticity and Cognitive Offloading in Evolving Robots.*, 2016.

ACM: Genetic and Evolutionary Computation Conference ACM Genetic and Evolutionary Computation Conference (GECCO), Berlin. Poster presentation: *Exploiting Environmental Differentiation to Promote Evolvability in Artificial Evolution.*, 2017.

ACM: SIGEVO Summer School SIGEVO Summer School on Genetic and Evolutionary Computation, Berlin. Participant, 2017.

IEEE: Symposium Series on Computational Intelligence 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu. Oral presentation: *Favoring the Evolution of Adaptive Robots Through Environmental Differentiation.*, 2016.

Signed: ____________________________

Date: ____________________________

Contents

List of Figures

Chapter 1

Introduction

Robotics has fascinated humans since long time ago. Leonardo da Vinci, for instance, designed complex automatons in the 15th century [113]. And other evidences point self-operated machines being thought and realized in China even before Christ [115]. Nowadays, robotics has been used in different domains as industrial and house environments [47], and even in the outer space [157]. While its use in a tele-operated way or in controlled environments, as in industry assembly lines, is very common, the use of mobile robots providing services in highly dynamic and uncertain environments is still a great challenge. This area, often called service robotics, has received much attention in the last years.

In the context of service robotics, uncertainty, noise and completely unknown environments are common working conditions, and dealing with all these things is a requirement. This means that adaptability is a key feature, the machines should be versatile in order to deal with these many different and unpredictable situations. The traditional techniques that allowed the development of many successful applications in industrial robots, strongly rely on accurate mathematical models describing the dynamics of the environment, the task and the robot. The appropriate use of these mathematical tools in the context of service robotics could be a very hard and even an unfeasible task. For this reason the traditional approaches for developing robot control are not sufficient anymore, at least not alone. Therefore becomes evident the necessity of new methods that could automatically embed robots with the proper adaptive behaviours for operating in this more dynamic and unpredictable scenario.

Indeed one of the possible definitions of adaptive behaviour for both animals and robots says that it is a behaviour that maintain the animal alive or that keep the robot functioning according to what it was designed for in these harsh conditions [109]. In more specific terms for robotics scenarios we could say that the term adaptive behaviour means: behaving in a given way through the time, according to the encountered environmental contexts, that will allow the agent to achieve a certain goal, given its sensorimotor capabilities and respecting a given set of conditions and constraints. Broadening the sense of this term, in the context of this thesis adaptive behaviour is used not only for referring to behaviours that enable the robot to achieve a desired function(s) in a given environment, but also behaviours organized in a way that facilitate the possibility to adapt to varying environmental conditions, and that facilitates the exploitation of integrated behavioural and cognitive capabilities.

It is important to notice that it is not possible to find in the literature a single, precise and widely accepted definition of adaptive behaviour. The definition of this term vary among the different scientific disciplines, and even inside a single discipline. In psychology for example, adaptive behaviour may be defined as "the collection of conceptual, social, and practical skills learned by people to enable them to function in their everyday lives." [182, p. 105]. A broader definition, usually used in biology and artificial life communities, says that adaptive behaviours are those that improve survival and reproduction rates of an individual [28, 74]. However it is noteworthy that the meaning of adaptation may vary even inside a same discipline as, for example, in biology when talking about evolution the meaning of adaptation is "any change in the structure or functioning of an organism that makes it better suited to its environment" [37, p.13] while in physiology it means "the alteration in the degree of sensitivity of a sense organ to suit conditions more extreme than normally encountered" [37, p.13]. Given that a precise and widely accepted definition of adaptation and adaptive behaviour is an open issue, providing a contribution to this discussion is not in the scope of this thesis. The definition adopted in the present work for is that presented in the previous paragraph.

An approach that is strongly related with adaptive behaviour is the embodied intelligence [26]. Differently from the traditional view of intelligence that works with symbolic representation in a perceive-process-act loop, embodied intelligence takes into account the fact that robots are situated and embodied agents that operate in a dynamic environment. The relation between the robotic controller - its 'brain' - the physical parts and the sensorimotor system - its body – and the environment is crucial for this way of viewing of intelligence. So more specifically, embodied intelligence [135] refers to the proper interactions among these elements that make the agent able to efficiently and adaptively perform the tasks. Finding the interactions that will produce effective behaviours, however, is not trivial for many reasons. For instance, manually programming the robotic behaviours supposes that the designer knows a priori the task solution, which for many complex tasks is not the case. Another challenge is the different perspective of the task with respect to the robot sensorimotor system and the human sensorimotor system. Perception and action are bounded through the environment which means that they mutually influence each other producing a loop, which is often called sensorimotor loop [133]. The regularities present in this loop have an important influence in which behavioural strategies an agent can perform for achieving a certain goal. The importance of these regularities is illustrated in many different works in the literature [8, 117, 124]. The strategies thought by a human tend to be limited by his understanding of which regularities may be present in the robot's sensorimotor system and in the robot/environment relation. This fact tends to exclude any behavioral strategy that could benefit of less evident regularities that are available to the robot.

Given these challenges, the use of adaptive methods, that enable the agents to develop their skills autonomously while they interact with the environment, represents a promising alternative for the synthesis of embodied and situated agents. These methods better explore the embodied intelligence that arises from the brain-body-environment interaction, usually not requiring manual programming of the robot behaviours. Indeed, the possibility to automatically discover the solution to a given problem, i.e. the necessary strategy for solving it and the manner in which such strategy is realized, enables to overcome the problems caused by the fact that human designers might be unable to identify appropriate solutions even on the basis of a detailed analysis of the problem domain [119]. For the moment however, the complexity of the problems that can be successfully tackled with these methodologies is still relatively limited. More specifically, the issue of whether and how adaptive methods can be scaled up to master "complex" problems, those requiring the development of a rich repertoire of integrated behavioural capabilities still represents an open challenge.

In the literature there are many works investigating the automatic generation of robot controllers. For achieving this goal, different techniques are used including Reinforcement Learning, Programming/Learning from Demonstration, Evolutionary Robotics and so on. These methods will be better presented in the next section as well as some works using these different techniques.

In this thesis I will focus on the study of the evolution of adaptive behaviours using Evolutionary Robotics techniques. These techniques use genetic algorithms for evolving robot controllers and/or body plans of situated and embodied agents in order to perform a given task [123]. Genetic algorithms [70] are a metaheuristic based on natural evolution that are often used for solving optimisation problems. Since this kind of algorithm usually operates without specific information about the problem dynamics, requiring only a way to evaluate the candidate solutions, it is often characterized as a black-box optimization technique.

Evolutionary Robotics though, might not be considered as a black-box optimization method [40]. As we will see in detail in the next chapter, there are some key factors that need to be defined by the designer and influence the outcome of the evolved solutions. Therefore, knowing better the underlying behavioural principles and mechanisms that could emerge from evolutionary methods may shed some light on how the designer should make decisions for creating processes capable of evolving more complex and effective behaviours without just simplifying the task or increasing the prior knowledge used about the possible solution itself. Indeed all the knowledge that may lead to a complexity increase of the learning tasks, and/or a reduction of the prior knowledge and the handcrafting required tend to represent an improvement on the evolutionary robotics field [116].

Evolutionary robotics methods have already presented interesting and effective behavioural solutions for different tasks as phototaxis [172], obstacle avoidance [148] and object categorization [14] among others. However, effectively using ER methods for synthesising robots capable of operating in real world applications is still an open challenge. One of the interesting aspects that have to be better understood to reach this milestone include the capability to display multiple behaviours and the behavioural selection that is adaptive to the current robot/environmental context. Another one is the capability to extract information from the environment and later use this information to regulate the behaviour of the agent. These two aspects represent two forms of adaptive behaviour, i.e. two forms of behaviour regulation.

In this context, this thesis will present studies that shed some light in the evolution of these two interesting types of adaptive behaviours: articulated and integrated cognitive ones. Despite the precise meaning of these terms will be clearer for the reader in the next chapters, here I provide a brief explanation of them. By articulated behaviour, I mean behaviour composed by sub-parts, or sub-behaviours with specialized functions, that will be combined together, in a sequential and context-dependent manner. This definition is inspired and aligned with the definition used by many biologists for characterizing animal behaviour [11, 53, 111, 184, 185]. It is noteworthy that it is possible to have a modular organization of the behaviour, without a modular and structural division of the robot controller. This will be clearer after the results presented in the third chapter of this thesis. By integrated cognitive behaviour I mean behavioural strategies that combine two different types of behaviour: (i) those relying solely on immediately perceived stimuli, therefore called reactive strategies, with (ii) those relying on information integrated over time and encoded into internal states that are used to regulate the robot's behaviour, therefore called cognitive strategies.

The second chapter of this thesis is intended to situate this work regarding the robotics and artificial life community, provide the reader with the proper background for fully appreciating the contributions provided by this thesis, and also to present the biological aspects that provided inspiration for the studies presented here. On the third chapter, I present a detailed study regarding the evolution of articulated behaviours for a single goal task and without the necessity of modular division of the robot controllers. The experimental setup include a single goal task, a cleaning task, in which a robot has to clean an unknown and varying environment. The evolutionary process is done with different robots that have different sensory capacities and different artificial neural networks architectures. The results obtained with this study indicate that it is indeed possible to evolve modular behaviour, for a single goal task, even without the explicit division of the robot controller in different modules. This type of articulated behaviour brings performance advantages compared to non-articulated ones, and its evolution and operation depends mainly on the robot's sensorimotor richness, which in turn will determine the robots capability to perceive and generate affordances, the main behavioural mechanism that allow the robot to produce long-term sub-behaviours and to properly alternate the behaviours in a smooth and timely manner. The affordance generation mechanism will be presented in details in the chapter 3.

In chapter 4, a study seeking to understand the role of reactive strategies in the evolution of cognitive solutions, i.e. those capable of integrating information over time encoding it on internal states that will regulate the robot's behaviour in the future, is presented. More specifically I tried to understand whether the existence of sub-optimal reactive strategies prevent the development of cognitive solutions, or they can promote the evolution of solutions capable of combining reactive strategies and the use of internal information for solving a response delayed task, the double t-maze. The experimental setup included again robots with different artificial neural network controllers that had to navigate in unknown and vayring environments in order to accomplish the task. The results obtained showed that reactive strategies capable of offloading part of the cognitive work to the agent/environmental relation can promote, rather than prevent the evolution of solutions relying on internal information. The analysis of these results clarified how these two mechanisms interact producing a hybrid superior and robust solution for the delayed response task. Finally the last chapter provides the general conclusions of this work, wrapping up and summarizing the main contributions regarding the evolution of these two types of adaptive behaviour.

Chapter 2

Background Knowledge

In this chapter the background knowledge regarding methods that automatically generate robot controllers and/or body plans will be presented, together with some of the works found in the literature using each of the methods. Furthermore, references for readers that could be interested in going down into the details of each them will be provided. Besides, this chapter introduces some of the biological knowledge regarding the evolution and use of adaptive behaviours in biological individuals, together with the different methodological approaches to obtain adaptive behaviour using evolutionary robotics methods found in the literature. Finally the last two sections of this chapter will situate the reader with a brief introduction specific to the two types of adaptive behaviours studied in this thesis in the remaining chapters: articulated and integrated cognitive behaviour.

2.1 Automatically generated controllers

Robotic automation has been used for many decades in the industry, and in this context robots are very common. We cannot state the same about mobile robots in the everyday life, i.e. robots operating in more common and less constrained environments like houses, offices, hospitals and so on. Commercially, only few robotic applications related to service robotics are available for naïve users. A well-known example of a commercial robotic application are the vacuum cleaner robots like, for instance, Roomba [75].

If we look at the industrial robotics scenario, we can see that they usually operate in a constrained and controlled environment, with very precise information about positions and velocities of objects and end effectors. In this industrial scenario, powerful mathematical tools are available, as for instance dynamical systems and control theory that enable the designer to model and properly control the robots according to the desired goals to be accomplished such as welding, moving objects from one point to another and so on.

However, in the context of service robotics the application of these methods has some severe limitations since it requires the knowledge of some aspects that are usually unknown. First, an accurate model of the robot and the environment should be created. This can be a hard task for both highly dynamic environments, like those in the everyday life, and complex robot morphologies. Even in cases when it is possible to acquire these models, it requires a great expertise from the designer, and since the more complex is the robot and the environment to be modeled, the more complex are the models to be created, this is also an error-prone task. Second, the sensory system of mobile robots provides only partial observations of the environment, and these often are continuous and noisy signals. So, in order to apply traditional control techniques, additional methods that estimate important variables, for instance positions of robots and objects, have to be used, increasing even more the complexity of these methods. Finally, the designer has to know a priori the solution to be achieved in terms of robotic behaviour, which for some tasks can be hard to know, if not impossible. A good example of unknown optimal strategy are the commercial vacuum cleaner robots that execute completely different cleaning strategies depending on the brand [1, 132].

An interesting alternative for the synthesis of controllers for mobile robots are methods that automatically generate the controllers using demonstrations or the robot's experience. These methods can be categorized in two groups: Learning from Demonstration and Trial and Error methods. It is still possible to subdivide the trial and error methods in two subgroups, the traditional reinforcement learning approach and its variations, which are formalized as a Markov Decision Process (MDP) [168], and Evolutionary Robotics methods [123]. Both subgroups rely on the robot's interaction with the environment in successive different trials. In this way, the methods progressively verify which actions produce the best results according to an objective function provided by the designer. This objective function is related with the task to be performed and usually does not require any information about the strategy to be performed by the robot. So these methods do not require neither demonstrations nor mathematical models of the robot and the environment.

2.1.1 Learning from Demonstration

In Learning from Demonstration (LfD) the controllers are synthesized based on a human teacher or a demonstration from another robot. In general, these methods are classified as a subset of supervised learning [5] since they receive predefined input-output mappings and the controller is generated in order to reproduce the previously received mappings. Many factors influence the complexity of producing controllers based on demonstrations. One of these aspects is, for instance, how the demonstration is provided, that is whether or not it is possible to create a direct mapping between the demonstration and the robot's sensorimotor system. Pure learning from demonstration methods have the requirement that the demonstrator has to know a priori the solution in terms of which would be the optimal behaviour for that task, which can be difficult for some tasks. Also the quality of the demonstrations, usually depending on the teacher proficiency when executing the task, has a strong influence on the success of the method. Since the demonstrator, often a human, has a very different perception of the world compared to the robot, it is hard for the human to see which solutions are feasible or not for the robot, or which solutions would be more or less effective considering the robot's perception capabilities. Methods that combine demonstrations and trial and error solutions remove this requirement, since the demonstration is used as an starting point and the final solution is a fine tuned version of it. Also the number of demonstrations in some cases can be a problem, since the fatigue factor has to be considered when using human demonstrators.

Many researches have been done using LfD to develop robot control policies. Lin et al. [98] proposed a framework to learn fingertip force for grasping and manipulation process from a human teacher with a force imaging approach. Konidaris et al. [89] proposed an online algorithm for constructing skill trees from demonstrated trajectories. Vakanski et al. [176] used demonstrations for performing the identification of key points in order to make the robot able to learn and to reproduce complex trajectories.

2.1.2 Reinforcement Learning

Traditionally, reinforcement learning techniques are formalized as a Markov Decision Process, which means that there exist a set of states, a set of actions, rules for transitioning between states and rules that reward being in a state and performing a certain action (sometimes only being on a state). So, the agent is let free to act in the environment having two main goals: explore and maximise the reward with its actions, which is often called the exploration-exploitation trade-off. This characteristic makes reinforcement learning a good option for online learning problems, where the agent has to execute some task and improve its execution at the same time. While this technique works very well when dealing with discrete, finite, and not too big state and action spaces, it suffers from the so called curse of dimensionality, where its applicability is strongly compromised when the state and action spaces are too big or continuous. Since in robotics, and even more in service robotics, high-dimensional, continuous, and usually noisy, state and action spaces are common, fitting the traditional discrete description of RL with this scenario in general does not produce good results [92]. For this reason, much effort has been done in order to create techniques that allow the use of RL techniques in the robotics field. The technique representing a great advance of the RL application in robotics are the policy-search methods, where instead of searching in a huge state/action space, a search is conduced in a much smaller space, the space of parametrized policies [92]. For detailed information about each of the different techniques involving RL see [85].

Some works that use reinforcement learning in robotics are [76] that uses brain activity of a human user observing a robot trying to solve a learning task as reward signals for the learning process. Konidaris et al. [88] proposed a method that combines reinforcement learning and decisions tree that enables a humanoid robot to learn a task with only few trials making the learning process faster than usual. Kormushev et al. [91] presented a reinforcement learning approach in order to minimise energy consumption of the walking of a passively-compliant bipedal robot.

2.1.3 Evolutionary Robotics

Another class of trial and error methods that automatically generate robot control policies are those involving Evolutionary Algorithms, this field is called Evolutionary Robotics [123]. Despite the main idea is similar to Reinforcement Learning, i.e. developing the control policy and/or the robot morphology through the robot's interactions with the environment, the mechanisms used in the evolutionary process are quite different. A very common evolutionary algorithm used consists in starting with a random population of individuals, representing candidate solutions, where each individual is encoded in a genotype. The genotype will be transcoded in the control policy and/or the robot morphology, and then evaluated. A fitness function defined by the designer is used for performing the evaluation and the individuals with greater fitness values tend to reproduce through mutation and/or recombination. This process is repeated for a given number of times defined by the designer or until some solution achieves a predefined fitness threshold. The idea behind the method is that the fitness function and the experimental conditions create selective pressures that will favour the survival of the fittest individuals, i.e. the individuals that best perform the learning task.

From a methodological point of view, at the end of the evolutionary process, if a satisfactory solution has not been found, the designer usually has two choices: abandoning the initial methodology and trying to develop the controllers in a different way. This was done, for instance, in [41], where a combination of handcrafted and evolved controllers was used. Or starting a trial and error process where the different parameters of the evolutionary method are changed to check which ones produce good results. Indeed, the less the knowledge of how the evolutionary process works and its underlying principles and mechanisms, the greater the designer blindness about the reasons why the process did not work well at first, how the changes will affect it, and which of the possible changes in the process are more likely to produce improvements. So, acquiring a greater knowledge about the fundamental aspects of ER techniques can lead to a better designer-orientated process. With this knowledge the designer can first investigate the possible reasons for the fail in the synthesis of an effective controller, and then, make the proper decisions about which aspects should be changed in order to make the development of effective controllers feasible.

The specific details of the different algorithms used in ER may vary. The genotypes, for instance, can be used for encoding the controllers in different ways, i.e. using different structures and techniques going from genetic programming [93] to fuzzy logic controllers [141]. A recent and interesting approach is presented in [149] where decision trees are used for evolving the controller of a flying robot. Currently, one of the most common setups within the field of Evolutionary Robotics is Neuroevolution [94], where the weights of an Artificial Neural Network (ANN) are defined by an evolutionary method. In the last years some works extended this approach, and instead of evolving only the ANN parameters, also the topology of the network is evolved as, for instance, in the NeuroEvolution of Augmenting Topologies(NEAT) [160] and Hyper-NEAT approaches [159]. While ANNs have the benefits of naturally dealing with continuous and non-linear dynamics, its combination with evolutionary techniques remove the requirements of training datasets used in supervised learning techniques. These aspects make the combination of ANNs and evolutionary computation a powerful tool.

The application of genetic algorithms to deterministic optimizations, for instance, in some engineering applications like [143], can be considered a black-box technique. This is mostly due to the fact that the fitness of each candidate solution is precisely obtained with each evaluation. In robotics, otherwise, the natural stochasticity present in this field has to be considered in the process. Due to the dynamic nature of environments, robots have to be able to deal with the performance of the learning task in different contexts, and to the fact that sensors and motors are noisy, it is not possible to perform a precise calculation of the individuals' fitness. In other words, there is intrinsic stochasticity in the evolutionary process which is an important difference compared to the application of evolutionary computation to deterministic problems. So, even with the execution of many different trials, it is only possible to estimate the fitness of each individual. In which contexts the robot will be tested, and how the environment varies along the different trials directly affect the generalization capacity of the robot to perform the task, i.e. the capability to solve conditions not encountered during the evolutionary process. Depending on the evolutionary algorithm used, testing the candidate solutions always in the same initial conditions tends to produce brittle solutions, i. e. solutions that do not generalise well to untested and real-operation conditions, therefore, varying the environment changing these conditions or adding random noise to them is important for obtaining robust solutions [67, 175]. On the other hand, huge variations among trials tend to prevent the process to converge to a good solution, since too abrupt changes would make harder to find individuals adapted to all the different conditions [50]. Finding the proper balance is not an easy task, and in general these parameters have to be adjusted by the designer according to the learning task.

Since artificial evolution tends to be a slow process that involves testing different individuals in millions of trials, apart from some simple tasks, in general it is unfeasible to perform all the evolutionary process using real robots. So, all or at least the first part of the process is often done through simulation. In order to successfully transfer the simulated evolved solutions to real robots it is necessary to cross what, in the literature, is called the reality gap problem. This problem arises from the fact that even with high quality simulators, those that have accurate models of the world, there will always be differences between simulated and real world. These differences prevent directly transferring the solutions found in simulation to real environments. In order to deal with this problem, other than having simulators that are accurate and as close to reality as possible, different approaches can be used. A first one is to model only the essential parts of the simulator and to create a sort of envelope-of-noise around it.

This can prevent that the simulated solutions rely on the specific simulated sensory-motor characteristics or simulated world dynamics, that could prevent their use in real robots. This approach was proposed and used in [79], and the simulated controllers were successfully transferred to reality. Another approach is to sample the sensors and motors activation on the real robot, and use these samples to fine tune the simulator in order to avoid discrepancies between simulated and real sensors and actuators, as done in [110]. In [90], simulated robots are evaluated not only based on how well they perform the learning task in silico, but also on how well they perform it in the real world. For doing so the individuals are occasionally tested in the real world, and based on that the simulator is adjusted to accurately approximate how transferable is each individual. Following a similar approach but going even further, in [140] the simulator used is completely created using real world data. The training sets are used to train an artificial neural network for simulating the sensory-motor flows experienced by the robot in the real world. The experiments presented in this thesis were performed only on simulation and some of the techniques presented above were used, as the addition of noise in the sensors and motors, and also the sensor sampling technique.

Another important aspect of ER is the fitness function used to evaluate the individuals. Although it is a scalar function, it involves many different aspects affecting the selective pressures applied over the individuals. Nelson et al. [116] categorise the different fitness functions based on the amount of prior knowledge added by the designer. According to the author they can be writen in terms of the behaviour that should be executed, the specific activation of sensors sets, or even in a binary way that says if a trial was successfully executed or not. Adding prior knowledge can be a good way to direct the search to a specific region of the search space. However, adding it could also be risky since the search can be biased and converge to a local optimum region, a problem usually called premature convergence or deception. Using less knowledge about the learning task, for instance using a binary evaluation of success/fail, means letting the process free to find the most effective behaviour that accomplishes the objective. However, depending on the complexity of the task, the initial candidate solutions could present an equally bad performance among individuals, and so a lack of fitness gradient would prevent the process to select the intermediary candidates necessary for finding the ultimate solution. This problem is usually called the bootstrap problem. So also in this aspect, the decisions made by the designer, which will directly influence the selective pressures, have a fundamental influence on the success of the evolutionary process.

Ollion et al. [40] present an interesting review about selective pressures in ER and some of the different aspects that exclude this field from the category of black-box optimization method. A different categorization of fitness functions is proposed, considering if they are oriented to the refinement of the goal to be achieved or to facilitate the exploration of the search space. They also propose a classification considering if the functions are task-specific or task-agnostic. Some other aspects and techniques related to Evolutionary Robotics used to tackle the deception and the bootstrap problems are also reviewed, like Novelty Search, Multi-objective Algorithms, Coevolution, Interactive Evolution, just to cite some of them. For detailed information about each of these aspects see [40].

Since there exist many different factors involved in the ER methods that require the designer's decision, a better understanding of the mechanisms and principles driving the evolutionary process can lead to the development of more effective and complex behaviours. This in turn, could make possible to scale up the complexity of the tasks tackled by ER methods. In this thesis, I will analyse the conditions that can promote the evolution of adaptive behaviours in artificial agents. More specifically those concerning the evolution of articulated behaviours, i.e. behaviours that are realized by combining different elementary behaviours, combined in sequence in a context-dependent manner. And also the conditions that promote the evolution of integrated behavioural and cognitive abilities. I believe that a better understanding of these aspects may enable the synthesis of the adaptive behaviours necessary for tackling real world problems. In the following section I briefly describe these aspects from a biological perspective.

2.2 On the evolution of adaptive behaviour

Natural organisms are capable of displaying adaptive behaviours that allow them to deal with different tasks and situations. The vast repertoire of behaviours found in natural organisms is the object of study of many different areas of biology as, for instance, ethology, that specifically studies animal behaviour. Additionally, and of particularly importance for the present work are those studies often found in evolutionary biology, that aim to understand the evolutionary origins of the animal behaviours, especially the complex ones.

Simple behaviour is considered to be a direct response of a stimulus, basically a reflex, while complex behaviours are those regulated by many external and internal factors [12, p. 27]. It is misleading to think that behavioural complexity is directly linked to the complexity of the individual's sensory-motor capabilities, as well as its neural capacity. It is important to notice that it is possible to display behaviours with a certain level of complexity even when the individual itself has limited capabilities. This is the case, for instance, of the navigation behaviour of desert ants [183]. For the case of artificial agents, for instance, intricate behaviours requiring the coordination of many different mobile parts as walking can be reproduced with simple neural controllers [35], or even with no active control at all [33].

A range of different factors can give origin to the dynamical processes we call adaptive behaviour. A first aspect is the existence of a rich environment that could offer stable dynamics, and stimuli that could be perceived and used by the agents in order to regulate their behaviour properly. In the case of the desert ant navigation behaviour, for instance, the polarized light patterns provided by skylight is used by the ant as a sort of compass. This mechanism allows the ants to perform a path integration strategy in order to estimate their position and perform a homing behaviour after the foraging behaviour was performed. Other than the path integration strategy, the ants are also able to use external landmarks, when available, in order to situate themselves in the environment and effectively navigate thousands of times the length of their own body [183].

Actually, the example of the passive walker [33] is another very interesting case in which the exploitation of the environment stable dynamics, the laws of physics, can provide an effective way of performing behaviour without any control. Of course, this is an extreme case considering the agent is completely passive. In this case, the external environment should provide not only the fundamental characteristics as, for instance, an inclined plane, but also the agent itself needs to be built with very precise and specific geometrical and material properties. With all these conditions met, the exploitation of the laws of physics for the production of the proper walking behaviour becomes possible. However, in some works the concept of the passive walker have been combined with active control, allowing the exploitation of the environmental characteristics and extending its use through the capacity of actively controlling the agent [128, 167].

The richness of the environment is an important aspect not only from the operational point of view, but also from the evolutionary one. Environments that provide different alternatives for achieving the same goal tend to facilitate the evolution of more elaborated behaviours, when those provide an advantage, compared to simpler environments. The example of the desert ants navigation can be used again if we consider that individuals capable of performing landmark-based navigation can be evolutionary selected only when the environment provide such landmarks. If the landscape is composed only by big flat portions of white sand with a blue sky, this kind of navigation will not be expressed and thus will not be selected. Even the presence of environmental noise can be exploited for the execution of effective behaviour. As in the case of an araneophagic assassin bug [186] that use noisy situations, for example the noise on the spider web caused by the wind, for moving through the web to reach its prey without being detected. The availability of certain types of resources can induce innovative behaviours. For instance, a given species of bird in Barbados started to demonstrate an opening sugar packets behaviour, due to the great availability of these in their region. While Birds from the same species but from other regions, with originally no sugar packets in it, did not perform the same kind of behaviour even when the packets were suddenly added in the environment [42]. A last good example of how the availability of certain stimuli influences the individual's evolution, in this case morphology, is eye degeneration in cave fishes. The eyes of a certain species of fishes that live inside caves, so with the absence of light, degenerate, clearly changing also its behavioural responses [192, 193].

A second important aspect regarding the evolution and use of adaptive and complex behaviours is the possibility to organise it in an articulated manner. Animal behaviours are often organised in functionally specialised subunits governed by switch and decision points [53]. Examples of elaborate behaviours including several different phases regulated through a rich set of context-dependent rules include the courtship behaviour of the grasshopper [129], the reproduction behaviour of female canaries [69], web construction and predation behaviours in spiders [44, 78]. Actually the possibility of composing behaviours in a modular way in animals can represent a simple way of building a vast behavioural repertoire as pointed by the evidences found, for instance, in studies involving frogs, owls, turtles and cats [18].

Indeed, the possibility to organise behaviour through the combination of different subunits brings advantages from the point of view of adaptiveness. Reusing and adjusting the smaller parts, that compose higher level behaviours, offer a flexibility that could bring advantages when dealing with different contexts and tasks, making the individuals more adapted to different situations.

The third important aspect regarding complex behaviours is the possibility of combining and using internal and external information for behavioural regulation. This aspect, which is also important in the case of articulated behaviours for subcomponents integration, switch and decision points, becomes fundamental when behaviours involving cognitive capabilities are required. The possibility of integrating information over time, enconding it into internal states, through conscious memory, physiologic states or even unconscious neural activation, is a primary aspect for the performance of cognitive behaviours, as for instance communication, reasoning and so on.

It is interesting to point out that cognitive behaviours should not be considered as exclusively internal processes, i.e., as processes based exclusively on internal or previously integrated information. Indeed, often the combination of internal stimuli with the perceived external ones define the individual behaviour in a coupled way. This coupled use of internal and external stimuli becomes clearer when we look at some evidences of studies with male dogs and rats, in which the spinal connections to the brain were lesioned, and still some copulatory responses happened normally when genital tactile stimulation was performed [12, p. 31]. Also the clear division between perception and action started to be challenged in the literature when the mirror neurons were discovered. This type of neurons tightly links perception and action by showing that there are brain regions that activate both when the individual performs a given action and when it recognises another individual performing it [54].

Regarding human cognitive behaviour, besides some evidences that the human brain use different sources of information for cognitive control, sensory stimuli, episodic, contextual memory and so on [86], the combination of internal and external behavioural strategies is very frequently used by humans. This can be seen, for instance, in tasks in which by performing a certain action, not directly related to the task, the load of internal, cognitive processing is reduced. For example the act of rotating the head for reading a rotated text, reducing the necessity of mentally rotating it [146]. Another case in which this kind of strategy is used is when a cook organises the ingredients of a recipe before start cooking, based on the order and moment each ingredient have to be used. This allows him to offload to the external environment, the necessity to remember the ingredients order [174]. In the literature this is often called cognitive offloading [145].

The decision of investigating the internal and external environmental aspects of behavioural evolution in robots was made for the following reasons: (i) the biological inspiration by itself that produces the hypothesis that some of the interesting behavioural outcomes observed in natural systems could be produced also in artificial ones; (ii) this direction is far less explored than others regarding the development of complex and articulated behaviours (this will be better discussed in the next section), and (iii) the interesting aspects observed in artificial evolution can also provide insights or hints to biological studies aimed to better understand animal behaviour and its evolution. As mentioned above, I also believe that the progress in the comprehension of these aspects will also improve our ability to tackle real life problems through the use of Evolutionary Robotics.

2.2.1 Different Ways of Producing Complex Behaviours for Artificial Agents

Behaviours regulated by different environmental factors, internal and external, mainly through an array of behavioural responses to different stimuli and environmental contexts, brings benefits compared with simple behaviours. This is not only confirmed by biological studies, but also by the fact that we can find a considerable amount of different studies in the literature seeking this goal of developing articulated behaviours in artificial agents [3, 6, 23, 24, 100, 103]. Although the goal is very clear, and there is no discussion on the fact that having a proper and rich behavioural repertoire provides a greater adaptability of the agents, how to develop this behavioural repertoire, which should be the different behaviours involved and how to regulate them is still an open challenge.

Many different approaches have been proposed for tackling these problems, an important and pioneering approach is called behaviour-based robotics [23], or the so called subsumption architecture. In this architecture adaptive behaviours are produced through the combined use of different simple and modular behaviours. Not only the results obtained with this approach are noteworthy, but it is also the influence it has had from then to the present day in the robotics field. Despite the subsumption architecture made a great contribution from the perspective of how to organise the robot controllers, some of the issues with this approach is that it leaves to the designer the complex task of properly dividing the overall behaviour into lower-level modules. It also requires from the designer the definition on how the arbitration among behaviours should be done. And finally since the modules are structurally separated in layers, reusing and adapting already existing behaviours for the creation of new behaviours usually means duplication, and so, the creation of redundant modules. It is important to notice that, as the subsumption architecture, many of the approaches for providing complex and articulated behaviours are based on a more engineered view where the designer usually needs to know and define a considerable portion of the architecture, models, modules and transition rules, which may not be feasible depending on the setup.

The combination of modular architectures with approaches capable of automatically generating the controllers, as for instance Evolutionary Robotics, is a promising alternative for tackling the problem of how to develop and arbitrate different robotic behaviours for a given task. Indeed, many researches can be found in this direction of developing modular behaviours for artificial agents, and there are many different ways of achieving that. One of them is through structural modular controllers, i.e., controllers composed by different modules that are responsible for the production of corresponding different behaviours that are grouped together and chosen through a specific arbitration mechanism usually operating in a context-based manner. In this category we could cite [41] in which different behaviours are implemented through different modules which could be manually programmed behaviours or evolved neural networks, and an arbitration mechanism that decided when to use each behaviour. Rahim et al. [142] evolved a neural network controller responsible for choosing the robot motor commands using as inputs the outputs produced by a set of pre-programmed modular controllers. In [22] through a specific control architecture, boolean networks, the author demonstrates how structural modularity can emerge spontaneously through an evolutionary method, and so, present multiple differentiated behaviours. Similar to the work presented in [22], in [73], through a modified version of the NEAT algorithm [160], the authors present an evolutionary technique capable of creating neural networks in which structural modularity emerges as a product of the evolutionary process, despite in this case there is no explicit arbitration mechanism. In [25], the authors present two ways of producing structural modularity which they call (i) fixed and hardwired, where multiple modules compete to control the robot, and (ii) dynamic, in which modules can be added through a new evolutionary operator, and a module-duplication operation is introduced in the evolutionary process. As conclusions this work shows how this new operator can facilitate functional specialisation through the development of multiple behaviours. The use of multiple alternative neural modules and a gating mechanism modulating the alternatives in order to determine the proper output, as the works presented in [25, 178, 195], is usually called neuromodulation.

Besides structural modularity and neuromodulation, some works seek to develop richer behavioural capabilities using artificial neural networks (ANN) associated with a mechanism called synaptic plasticity. In the context of ANN's the term neural plasticity refers to the network capacity to dynamically change its synaptic connections, or weights, making possible the execution of differentiated behaviours over time. Martin et al. [104] show how the use of synaptic plasticity can allow the emergence of complex locomotion patterns for a exploration task. In [63], besides a modular neural controller, synaptic plasticity is used for creating a walking behaviour that can adapt to different contexts, making it able to avoid obstacles and navigate in complex environments. Steingrube et al. [164] showed how through the use of synaptic plasticity, a simple neural network could be able to adapt and exhibit many different behaviours, as taxis, self-protection, different gaits pattern and so on, even in a complex robotic setup of eighteen sensors and eighteen motors. Synaptic plasticity is indeed an interesting way of producing effective behaviours and Floreano and Urzelai [52] present explicitly this idea, arguing and showing some evidences about how encoding rules of synaptic plasticity in the evolutionary process could contribute to the evolution of more complex and adaptive behaviours.

Alternatively to the use of hierarchical and structural changes in the robot's controllers, specifically regarding artificial evolution methods, there are different approaches that seek to promote changes in the evolutionary algorithm as the main drive to obtain articulated and mainly more effective behaviours. In order to illustrate this kind of approach, I will present here three different kinds of approaches that focus on the evolutionary algorithms themselves, Evolution by Viability [102], Coevolutionary Approaches [121, 138, 161] and Novelty Search [96]. From these, the one that is currently receiving much attention in the literature is novelty search [36, 95, 96, 179]. Proposed in [96] the rationale behind this approach is that natural evolution does not work with specific objectives, as the specific goal functions commonly used in artificial evolution. So instead of evaluating individuals with an objective metric, the evolutionary selection is performed based on a behavioural novelty metric that considers all the behaviours found in the process. The main argument for doing this is that looking directly at the goal can cause the artificial evolution to be stuck in local optima regions, the so called deception or premature convergence problem. While the results presented in [95, 96] confirm that novelty search in some cases can reduce the deception problem, allowing the evolution of more effective behaviours, the development of the behavioural metrics that are used to define novelty is not a simple task at all. Also as the task complexity increases, developing behavioural metrics that produce effective behaviours can become as hard or even harder than using objective fitness functions. Furthermore, from the point of view of the inspiration by natural evolution, it is totally unrealistic keeping a record of all behaviours already seen in the evolution with the purpose of selecting new behaviours, despite this is not a central issue.

Another approach in this category that has proven to present effective results is coevolution [121]. In this approach two or more populations are evolved together and their fitness' depend on one another. The way in which the individuals interact defines if the process is cooperative, i.e., the individuals are composed together and cooperate to increase their fitness [138]. Or competitive, meaning that individuals compete, so usually a fitness increase of a given individual means a fitness decrease of the competing individual [51, 161]. A third approach called Evolution by Viability [102] focuses on the evolutionary selection, it operates by selecting all individuals that are considered viable based on a specific range of multiple fitness functions. The fitness range used to define if a individual is viable or not is progressively adapted based on the proficiency of the whole population. The viable fitness range at the beginning is wider, and when individuals start to become effective on performing the task, the range starts to become more and more constrained, gradually putting more selective pressure on the individuals.

Additionally, some noteworthy evolutionary approaches are open-ended evolution [158] and incremental approaches [163]. Open-ended evolution approaches take the biological inspiration on natural evolution one step further, considering that on natural systems we have an unconstrained reality. This in turn results on individuals and populations that keep evolving through the time producing increased complexity as generations go on. So the main idea on this approach is approximating artificial evolution to the natural one, removing the task-driven objective fitness function, and letting the selection to be done by the environment in an open world in which its dynamics dictate which individuals will survive and reproduce. Besides the interesting results obtained with non-robotics simulators often used in the literature, as Tierra and Avida [158], many works try to develop robot bodies and behaviours using open-ended evolution. For example, Bianco and Nolfi [17] showed how interesting behaviours as chasing and evading from other individuals or collective obstacle avoidance can emerge through the use of elementary robotics units and open-ended evolution. In [97] a novel method called ESP (encapsulation, syllabus and pandemonium) is proposed for creating behavioural and body complexity in evolved virtual creatures using open-ended evolution. Despite open-ended evolution is a promising approach it is not simple to use it for practical applications since in theory this kind of method is not directed to a specific goal. For this reason some approaches as that presented in [126] try to combine open-ended and task-driven evolution for obtaining evolved robots capable of solving specific tasks.

Incremental approaches try to reproduce the same incremental individual complexity produced by natural evolution in a more straightforward way. Therefore, incremental evolution approaches use task complexification for trying to push the evolutionary process to produce increasingly complex solutions. This can be done through the functional or environmental complexification of the evolutionary process [10]. As presented in [10] a functional complexification means increasing the rigor with which the individuals are evaluated along the evolution, relaxed evaluations at the beginning and demanding ones toward the end. Environmental complexification means increasing the complexity of the environment and the environmental conditions along the evolution, simple cases and tasks at the beginning and hard cases toward the end. For example, Stanton and Channon [162] showed how through the environmental complexification of the task, dividing it in different levels of difficulty and incrementally using them along the evolution, it is possible to evolve virtual agents in a 3D world to solve a rather complex task using artificial neural networks. How to divide the task, and how to present the subdivision of the task to the population is an important challenge to be tackled for using incremental evolution. In [112] they use incremental evolution for creating virtual creatures capable of crossing a wall of different heights, they provide an interesting analysis on how the possible ways of dividing and presenting the environments to the creatures during the evolutionary process can significantly change the outcome of the artificial evolution. And in [45] the authors propose a method that automates the reconfiguration of the test environment, allowing the use of incremental evolution in real robots without the necessity of human intervention.

As we could see, changes in the control structures, the use of hierarchical or modular structures, or in the evolutionary algorithm itself, i.e. changes in the fitness function or in the evolutionary selection, can contribute to the production of complex and effective behaviours. Indeed, besides the examples presented in this thesis, a great variety of approaches can be found in the literature that fits in one of these categories. However, in the context of Evolutionary Robotics, the investigation of potentially important artificial evolution aspects, that are present on natural evolution and biological systems, as for instance the evolution of articulated and cognitive behaviour without the use of specific algorithms or additional mechanisms introduced by the designer, is not a so common direction taken. This is a promising direction for some reasons, first of all

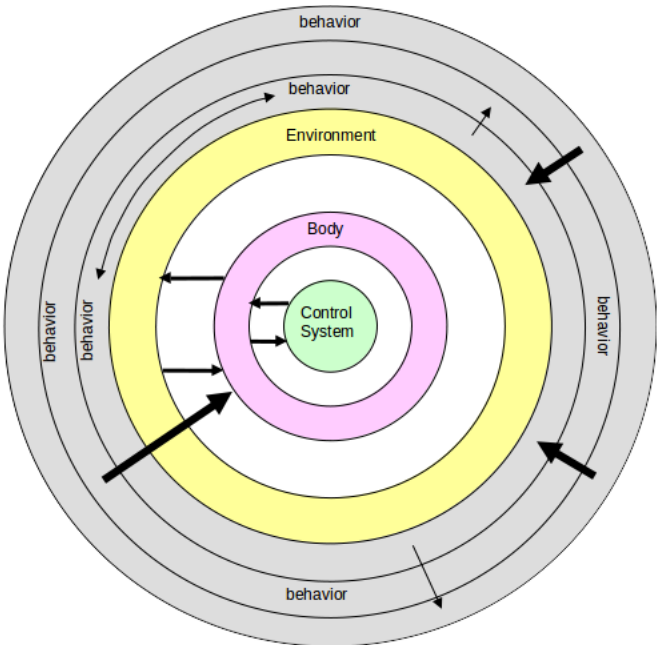


Figure 2.1: The overall behavioural architecture of evolved embodied agents showing the organisation of articulated behaviour as a multi-level structure achieved through the entangled combination and the interaction among the elements as presented in [120]. Adapted from [120].

the fact it has not been extensively explored yet leaves room for the study of many different aspects that are present in biology and that could apply to artificial evolution as well leading to improved results when evolving adaptive behaviour for robots. Another aspect is the fact that, in the long term, the insights provided by artificial evolution could help to shed some light, or at least to provide an additional tool for the study of not well understood aspects of natural evolution, mainly regarding the evolution of adaptive behaviour.

Actually the biological view of complex and articulated behaviours is well aligned with the fields of embodied intelligence and evolutionary robotics. In these research fields, behaviour is seen as the dynamic process that emerges from the interactions between mind, the robot's controller, its body and the external and internal environment [120, 135]. Furthermore, the possibility to build behavioural complexity building up on top and reusing previously acquired behavioural capabilities, see figure 1.1, can lead to the evolution of interesting and effective solutions as shown in [120].

Considering the similarities regarding behaviour found in biological and evolutionary robotics researches, drawing inspiration from biology in order to develop more effective articulated behaviours in artificial agents seems to be a promising direction. Indeed, natural evolution was capable of evolving biological creatures with amazing behavioural capabilities and that are capable of dealing with different tasks and situations. Considering that the ultimate goal of service robotics is to develop robots capable of operating in complex, dynamic and often unknown environments, for providing specific services, the biological inspiration might provide valuable insights for dealing with this open challenge. In other words, different fields of biology as, for instance, evolutionary, developmental and ethology, could point out interesting directions on how to tackle this current open challenge: how to develop intelligent robots capable of operating in the different environments of everyday life.

We know from ethology that the animal's environment has a huge influence on its behaviour [43]. Also from evolutionary biology we know that behaviour is a phenotypic response of an organism [185], which means that it is co-determined by the individuals genotype and the environment both internal, i. e. the characteristics and dynamics of the individual itself, and external, the world in which the individual lives. As already presented, these biological aspects are well aligned with the embodied intelligence theory [135], which argues that adaptive behaviour arises from the interactions among the robot's body, mind, a.k.a its controllers, and the environment.

The analysis and experiments described in this thesis investigate to what extent these aspects, often studied in the context of animal behaviour, are relevant for the evolution of artificial agents displaying adaptive behaviour. More specifically, the first one regards the evolution and fundamental aspects of behavioural plasticity in artificial agents. This term refers to the individual's capacity to present an array of behavioural responses that will be used based on internal and external stimuli without the necessity of structural controller division or synaptic plasticity. The second aspect concerns the role of cognitive offloading, i.e., reactive strategies that can extend the cognitive capabilities of an agent by offloading part of the cognitive work to the external environment, in the evolution and operation of cognitive robots. These topics will be briefly presented in the next sections, and the experiments and results regarding each of them will be presented in detail in chapters 3 and 4.

2.2.2 The Evolution of Behavioral Plasticity

Phenotypic plasticity is a term that refers to the capacity of a genotype to be expressed in different ways regarding its external and internal environmental conditions, i.e. in response to different environments. Some of these changes are irreversible, as for instance the sex determination driven by the environment in some reptile species [80]. Others are seasonally reversible [154] as for instance the colour of some species of insects that change according to the period of the year, autumn or summer [68]. Finally, behaviour is a special case of phenotypic plasticity that is reversible and occurs at a faster pace when compared with morphological plasticity [185, p. 27]. The main point regarding phenotypic plasticity is that it could be an important feature for the evolution of adaptive individuals able to deal with heterogeneous environments or dynamically varying environments [39]. A study that can illustrate this benefit in nature is [64], which presents evidences that the behavioural adjustment of a specific species of birds regarding singing in a anthropogenically noisy environment is done by short-term adjustments instead of evolutionary adaptations, i.e. they adapt based on behavioural plasticity. It is important to notice that behavioural plasticity should not be mistaken by the developmental process that involves the adjustment of synaptic connections among neurons, called synaptic plasticity.

Linking the concept of phenotypic plasticity to the evolution of intelligent robots seems to provide an interesting direction, specifically concerning the development of behavioural plasticity in evolving robots. An individual's genotype capable of presenting different behaviours according to the environment in which it operates is well aligned with the goals pursued by roboticists, especially with respect to mobile robotics. So, robots able to present articulated behaviours, those composed by different alternative sub-units regulated on the basis of internal and external environmental information, may lead to an increased performance compared to robots performing uniform, i.e. non-plastic, behaviours.

As previously presented, articulated behaviours in artificial agents have been developed through different approaches. In this work however the objective is to check the possibility to evolve behavioural plasticity without the use of structural modularity, synaptic plasticity or specific algorithms that could facilitate the emergence of modular behaviour. In chapter 3, this matter will be investigated in detail, showing how behavioural plasticity can emerge in a single goal task, bringing performance advantages when compared to non-plastic behavioural strategies. Furthermore, a detailed analysis of the emergent modular behaviour will show the importance of connectedness and relatedness among behaviours, and also how the possibility to generate opportunity for actions, i.e., affordances [55], is an important underlying mechanism allowing the effective evolution and use of this kind of behavioural strategy.

2.2.3 Cognitive offloading and the Evolution of Agents with Cognitive Capabilities

The individual's ability to display adaptive behaviour is increased if it is able to present cognitive capabilities rather than only direct responses to environmental stimuli, i.e. completely reactive behaviours. Since it is possible to find different meanings for the term cognition in the literature, here the term is used in a wider sense. More specifically, it means the capacity to integrate sensory-motor information over time, modifying so the individual's internal environment, and using the integrated information for behaviour regulation. The combination of reactive behavioural strategies, i.e. those relying on the immediate perception of external stimuli, with those relying on internal information integrated over time, seems to be another important aspect in the evolution of complex behaviours.

An important concept that lies at the interface between reactive behaviour and cognitive capacities is cognitive offloading in which the information used to regulate the agent's behaviour is not maintained in the agent's internal states but rather in the agent/environmental relation (i.e. it is offloaded in the environment). It is common that these actions and strategies might not be directly related to the goal task, but they have a fundamental role in the main cognitive strategy. As the term suggests this kind of action partially offloads to the external environment the cognitive workload required for solving a given task. As examples of cognitive offloading we could cite ordering the ingredients of a recipe before starting to cook, tilting our heads when trying to understand a given image or even using paper and pen for solving mathematical equations. There are evidences that show how offloading part of the cognitive process can extend the cognitive capabilities of an individual, minimise efforts, and improve performance in different domains as perception, memory, spatial reasoning and so on [145].

Specifically in artificial agents, cognitive offloading usually means acting in a way so as to encode information in the external environment or in the agent/environment relation in a way that it could be readily used for behavioural regulation. The use of cognitive offloading strategies is often seen in reactive agents, those with controllers that are structurally unable to integrate information over time, and so the only possibility to overcome their limitations is adopting these strategies. To illustrate some examples of this kind of strategy we could cite [16], in which a minimally cognitive agent capable of catching objects with a given size while avoiding others in a categorisation task used information offloading in order to properly perform the task. In this task, objects moved vertically, falling down along the trial while the agent moved horizontally. The behavioural solution find by the agent was travelling a certain distance proportionally to the object size, and after not perceiving the object anymore it started to go back. By behaving this way it stored the object size in its relative position to the environment which guaranteed that it would catch only objects of a specific size while avoiding the others. In [29] an evolved reactive agent, i.e. with no capacity to integrate information over time, capable of placing and perceiving environmental markers becomes able to perform a homing task through the use of a sort of breadcrumb navigation. In [118] a khepera robot, with a simple reactive controller, becomes able to differentiate ambiguous stimuli when perceiving a food object and the walls. The robot developed a behavioural pattern that explored the environmental regularities, and doing so its behaviour converged to a limit cycle that allowed the agent to remain close to the food objects, but not the walls.

Although it is well known that this kind of strategy usually emerges for overcoming individual's cognitive limitations, as the ones presented before, it is not clear what is the relation between the development and use of cognitive offloading and internally cognitive strategies. Actually, according to some authors, the development of reactive solutions blocks the development of cognitive capabilities. In particular, [127] claimed that reactive solutions constitute hard to escape local optima that prevent the development of cognitive solutions. Similarly, [95] claimed that the deception caused by the availability of locally optimal reactive policies is one of the main factors that explains why it is difficult for cognitive policies to evolve. For these reasons the authors hypothesised that the development of cognitive solutions necessarily requires specific selective cognitive pressures such as fitness components that encourage the development of short-term memory [127] or mechanisms for avoiding local optima, such as novelty search [95].

In a broader sense, we could say that the exploitation of the information that can be extracted directly from the environment and the effects of situated actions do not only affect the agents' low-level capabilities. Embodied and embedded strategies co-exist and interact with different strategies that are less dependent on agent/environmental interactions and more dependent on internal processes at all levels of organisation [31]. However, the relation and the interaction among strategies and capabilities that differ in that respect have not yet been investigated. Consequently, the question of how these different types of strategies can be integrated from an operational and developmental perspective is still open. In particular, one important question that needs to be answered is the following: "Is cognition truly seamless – implying a gentle, incremental trajectory linking fully embodied responsiveness to abstract thought and off-line reason? Or is it a patchwork quilt, with jumps and discontinuities and with very different kinds of processing and representations serving different needs?" [31].

The present work, specifically in chapter 4, investigates the relation between the development of reactive and cognitive capabilities. In particular, the objective is to investigate whether the development of reactive capabilities, specifically cognitive offloading strategies, prevents or promotes the development of cognitive capabilities, and how do they interact. For the scope of this work, cognition is defined as the ability to integrate sensory-motor information over time into internal states, and to use these internal states to regulate the way the agent reacts to perceived stimuli. The term cognition is often used in a more restricted way. The above definition is focused on a fundamental capacity being at the basis of all cognitive capabilities (e.g. perception, memory, attention, decision-making, reasoning, language, etc.). So in chapter 4 the results showing the importance and the synergy provided by the early development of cognitive offloading to the posterior development and use of cognitive strategies, that make use of both external and internal information, will be presented in detail.

Chapter 3

Behavioral Plasticity in Evolving Robots

As discussed in the previous chapters, an important aspect of adaptive behaviour is modular organisation, i.e. the fact that behaviours can be organised into a set of semi-independent sub-behaviours that are combined in a context dependent manner to achieve the agent's goals. Behaviour therefore can be organised into partially independent elements that can be recombined in a modular way even if they do not necessarily correspond to different modules or parts of the agent's nervous system. The utilisation, in a context-dependent way, of behaviours organised in this manner is indicated by biologists with the term behavioural plasticity. In this chapter we will see how behavioural plasticity can emerge in evolving robots and the advantages it provides.

More specifically, this chapter aims to study the possibility to evolve articulated behaviours even for a single goal task and without the necessity to divide the robot's nervous system into different modules. The experimental setup used is a cleaning task, which has a single goal, where robots with different sensorimotor capabilities and artificial neural networks architectures had to clean an unknown and changing over time environment. The results confirmed the possibility to evolve this type of adaptive behaviour even for a single goal task, and also that this behavioural plastic strategy can bring performance advantages compared to non-plastic ones. Analysis of robots' behaviours helped to clarify that the evolution of this type of behaviour depended on the characteristics of the neural network controllers and the robot's sensorimotor capacities, that in turn determined the robot's capacity to perceive and generate opportunity for actions, which in psychological literature is often called affordances.

3.1 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" [185, 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 [87].

This kind of plasticity can be seen as a key aspect of animal behaviour since it 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, from the point of view of the adapting individuals, what matters 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). Since moving around an environment with different characteristics will produce different perceptual stimuli to the agents, this means that environments are mostly variable from the perspective of an organism that is situated and performs actions in it, independently of whether or not they appear variable from the perspective of an external observer.

A relevant concept for studying behaviour, especially in the context of this study, is affordance. The term was coined by a psychologist called Gibson [55], and it provided an important and alternative direction for studying behaviour and cognition. According to Gibson [55] objects and the environment are perceived by an individual according to which actions they do afford, i.e. opportunities for action. Furthermore, these affordances are not mental representations that require cognitive processing for being perceived, but they are directly perceived from the environmental stimuli. This alternative way of viewing behaviour and cognition, which strongly depends on the environment and the individual and for this reason is often called ecological, challenged the traditional computational model of the mind, in which mental representations are built using the external stimuli and all the processing and action selection is done over the mental representations. In the last years, many studies in different areas focused on the study of the affordance concept and the creation of computational models for it [171]. The ecological view of affordances theory is well aligned with developmental robotics and embodied intelligence theories and for this reason many works on these fields are focusing on the study of affordances as well. For a extensive and recent review on the different interpretations, models, and the use of affordances in the robotics field, see [194].

This chapter presents a study that analyses 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 behavioural 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.

3.1.1 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 embodied and situated agents, 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 (imagine, for instance, a body moving with different friction values) 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.

The overall behaviour might be characterised by a modular organisation with somewhat semi-discrete and semi-dissociable subunits [185], 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 [119]). 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 of 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 [185]. 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 that are not only capable of displaying behaviours characterised by a modular organisation, but are also capable of displaying the capability to properly 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 those described above and to the capability of selecting 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 [122].

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 [11, 53, 111, 184, 185], others consider behavioural units as useful fictions at best [46]. 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 [23] 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 [82, 119, 139]. The attempt to resolve this issue is outside the scope of this study. 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 [185, 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 recognise 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).

3.1.2 Related works

Evolutionary robotics [120, 123] concerns the synthesis of a 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 [49, 122, 173]. Nolfi and Parisi [125], 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 [51] 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. This is an effective way of acquiring some degree of behavioural plasticity, i.e. adjusting an individual's behaviour - the way it responds to a given set of stimuli - according to changes in the environment.

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 [25, 151]. 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 interconnections [22, 73, 180].

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 [152]. 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., [60, 142]). In other cases, however, the behaviours were evolved as well [77, 134, 153, 189]. 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 [134, 153] or moving on the basis of a wheeled or legged actuators [189]].

Aiming 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 a series of experiments was performed. As we will see, the 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 [27, 55]. 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, the richness of the robot's sensory–motor system, and 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-behaviours and the importance of providing the agents with mechanisms that enable them to realise a smooth and effective transition between sub-behaviours. Details about the experiments, the results and detailed analysis' are presented in the next sections.

3.2 The method

To study this issue, we decided to consider a cleaning experimental scenario in which a wheeled robot needs to vacuum-clean the floor of an unknown in-door environment. We chose 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 [75]). 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 portion 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 chose 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 [1] demonstrates that finding the optimal solution/s for this problem is far from trivial.

3.2.1 The task, the environment and the robot

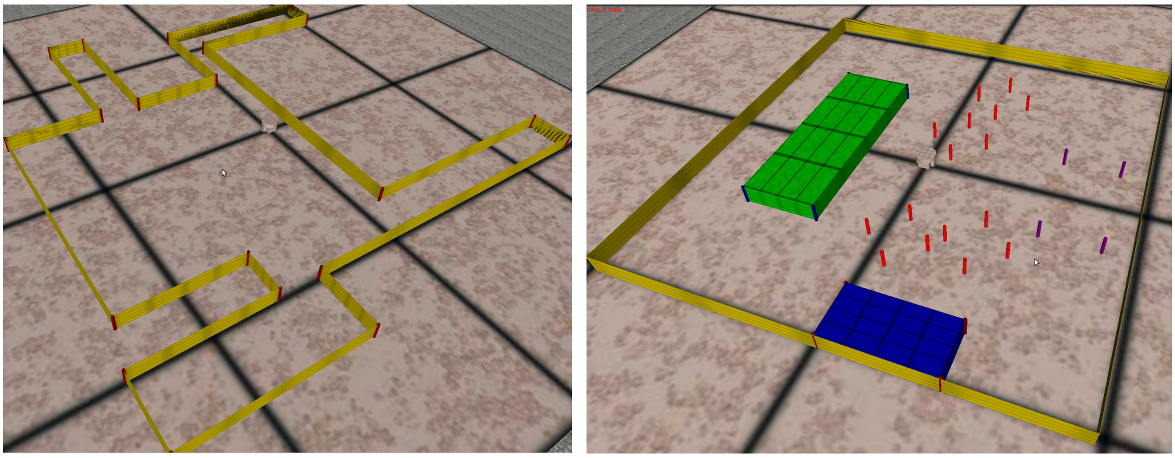
To evolve robots that are robust with respect to environmental variations, we evaluated each robot for three trials or cleaning sessions. At the beginning of each trial, the initial position and orientation of the robot in the environment, and the specific characteristics of the environment, like dimensions 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 robots to clean completely or almost completely a surface with similar properties is impossible due to the lack of data (for some indications see [132]), this represents a rather short period of time.

To compute the cleaning performance, we divided the environment in 20×20cm non-overlapping cells, and calculated the number of cells visited by the robot at least once during a trial. Equation 3.1 presents the fitness function where T_{tn} is the total number of cells in each of the three trials, which is different for each trial since the environment dimensions vary among the different trials. However, the same three trials are used for all the individuals of the same generation.

$$fitness = \sum_{tn=1}^{3} \frac{1}{3} \sum_{i=1}^{T_{tn}} C_i, \; where \; C_i = \begin{cases} 1 & \text{if } C_i \text{ was visited by the robot} \\ 0 & \text{if } C_i \text{ was not visited by the robot} \end{cases} \tag{3.1}$$

The experiments have been repeated in two different types of environments. In the first set of experiments, we used a concave environment (Fig. 3.1, left) constituted by a large central area and by four peripheral corridors that represented a room-like environment. The average environment had a central area with a size of 6.8m^2 and four corridors with a size of 3.78m^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 experiments, we used a convex environment



(a) (b)

Figure 3.1: Examples of concave and convex environments, (a) and (b), respectively

(Fig. 3.1, right) constituted by a rectangular room-like area including furniture. The rectangular area had a size of 12.2m^2±33% and included: a first rectangular object with an area of 0.93m^2±10%, a second rectangular object with an area of 0.17m^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 [20], a differential drive wheeled robot with a diameter of 17cm. It was equipped with 24 infrared sensors evenly distributed along the robot's body and capable of detecting objects in a range of 10cm. Moreover, it was also equipped with a rotating laser sensor capable of detecting obstacles at longer distance (1m). Experiments were run in simulation using the FARSA open-software tool [106, 107] that includes an accurate simulator of the robot and the environment.

3.2.2 The robots' neural controller

The robots were provided with a neural network controller. In all experiments, the robots were 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 encoded the desired speed of the two robot's wheels. The sensory neurons were fully connected with the motor neurons and to hidden neurons (if present), and the

Figure 3.2: Simple Neural Network Controller with (b) and without (a) hidden neurons. The output neurons W define the motor activation for left and right wheels of the robot. Neurons with a white contour around it mean the presence of a bias parameter.

hidden neurons were fully connected to the motor neurons. Hidden and motor neurons were provided with biases. The states of the hidden and motor neurons were computed on the basis of the logistic function. The state of the sensory neurons and the desired speed of the robot's wheels were updated every 50ms. Experiments have been replicated in the following four experimental conditions:

(S) Simple: The robots were only provided with the infrared sensors (see Fig. 3.2)

(R) Range sensor: The robots were provided with an additional sensory neuron that encoded the average distance of obstacles located within 1m 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 (see Fig. 3.3)

(T) Time: The robots were provided with an additional sensory neuron that encoded 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 enabled the robot to access information

(a) (b)

Figure 3.3: Range Neural Network Controller with (b) and without (a) hidden neurons. The output neurons W define the motor activation for left and right wheels of the robot. Neurons with a white contour around it mean the presence of a bias parameter. The input AvgL is the average activation of the rotating long range sensor.

extracted from the robot's internal environment (e.g., a robot clock situated inside the robot body), while the other sensors enabled the robot to access information extracted from the external environment (see Fig. 3.4)

(M) Modular: The neural controller was formed by three modules (each provided with eight infrared sensors connected to the two motor neurons) that were 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 (see Fig. 3.5)

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 robots were 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, were 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

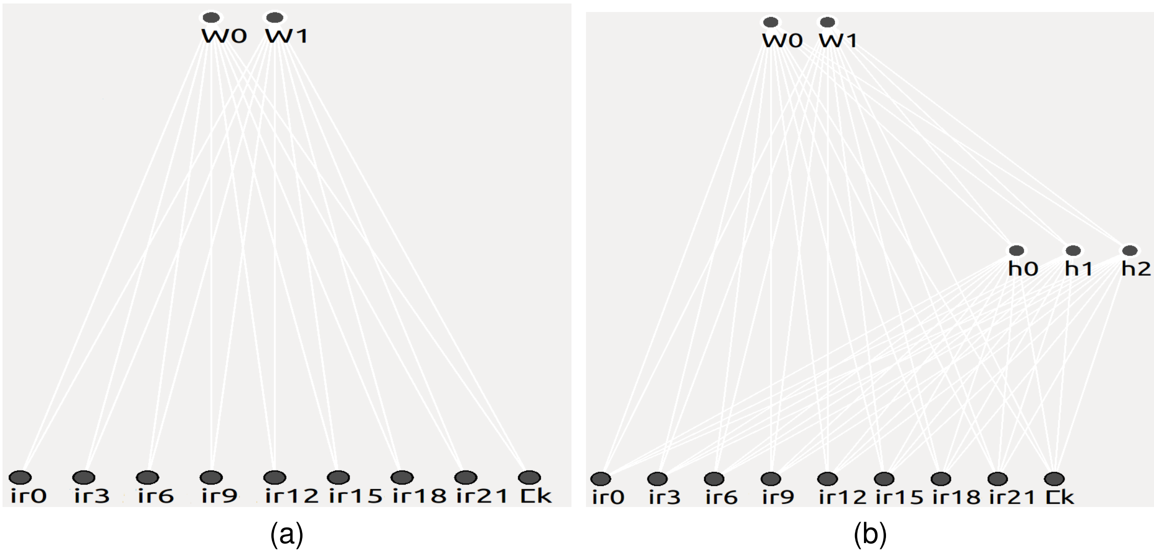


Figure 3.4: Time Neural Network Controller with (b) and without (a) hidden neurons. The output neurons W define the motor activation for left and right wheels of the robot. Neurons with a white contour around it mean the presence of a bias parameter. The input Ck is the internal clock that varies linearly from 1 to 0 along the trial.

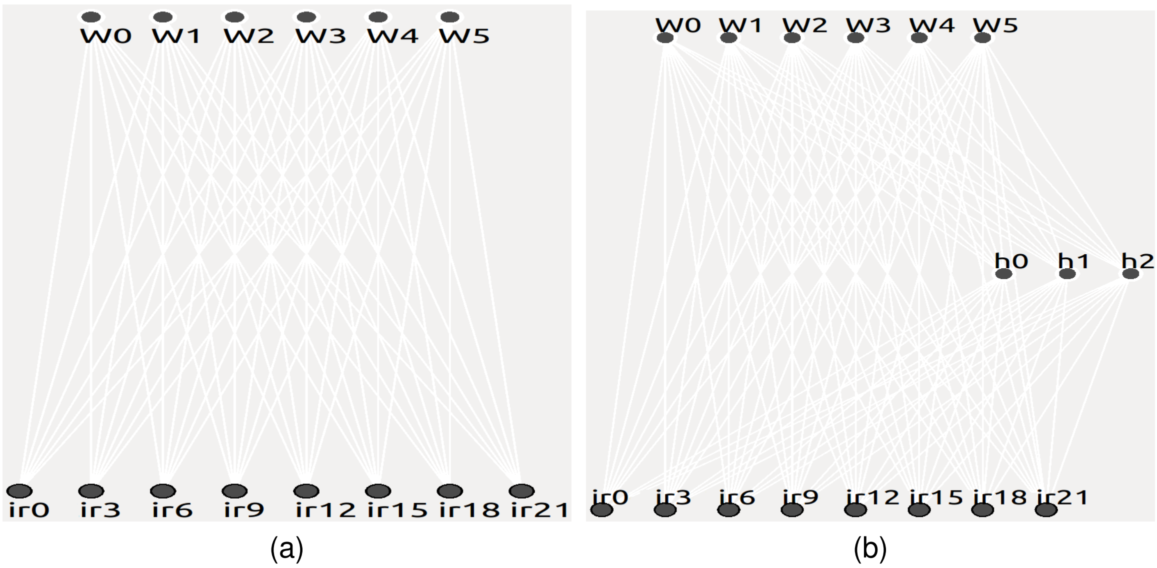


Figure 3.5: Modular Neural Network Controller with (b) and without (a) hidden neurons. The output neurons W define the motor activation for left and right wheels of the robot. Neurons with a white contour around it mean the presence of a bias parameter. In the first third of the trial output neurons W0 and W1 define the motor output of the robot, in the second third neurons W2 and W3 define the motor output and in the last part of the trial neurons W4 and W5 define the motor output.

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 [125]). However, all these experiments led to poorer results with respect to the base (M) condition.

3.2.3 The evolutionary algorithm

The initial population consisted of 20 randomly generated genotypes, which encoded the connection weights and biases of 20 corresponding individual robots (each parameter was encoded by 8 bits and normalised in the range [–5.0, +5.0]). Every generation, each individual was evaluated for three trials in environments that randomly varied in dimension within the limits indicated above. The fitness of each trial was calculated by counting the number of 20×20 cm portions of the environment that were visited by the robot at least once during the trial. The total fitness was calculated by averaging the fitness obtained during the three trials, see equation 3.1. All individuals were allowed to generate an offspring that was also evaluated for three trials. The 20 offspring were generated by creating a copy of the parent genotype and by mutating each bit with a 2% probability. The genotype of the offspring was used to replace the genotype of the worst parents or discarded depending on whether or not offspring outperformed 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 and the evolution lasted for 500 generations in each replication.

3.3 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.

3.3.1 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. 3.6, 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

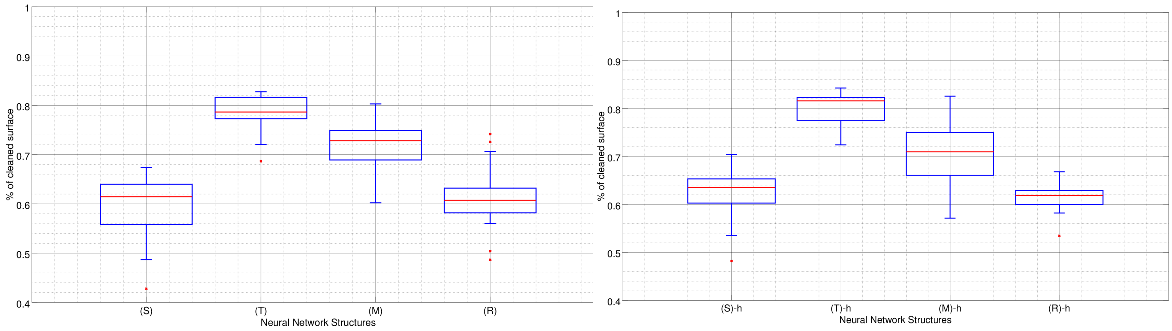


Figure 3.6: Box plots of performance in the concave environment. The left and right figures report the results obtained without internal neurons and with internal neurons, respectively. The box plots display the performance of the best robot of the last generation in the four experimental conditions, 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

performance obtained in the experiments in which the robots were also provided with the internal neurons (Fig. 3.6, 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. 3.7, top).

Indeed, the best robot with a simple architecture (S) always behaves in the same manner during the successive phases of the trial (Fig. 3.7, top-left). In particular, it avoids 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

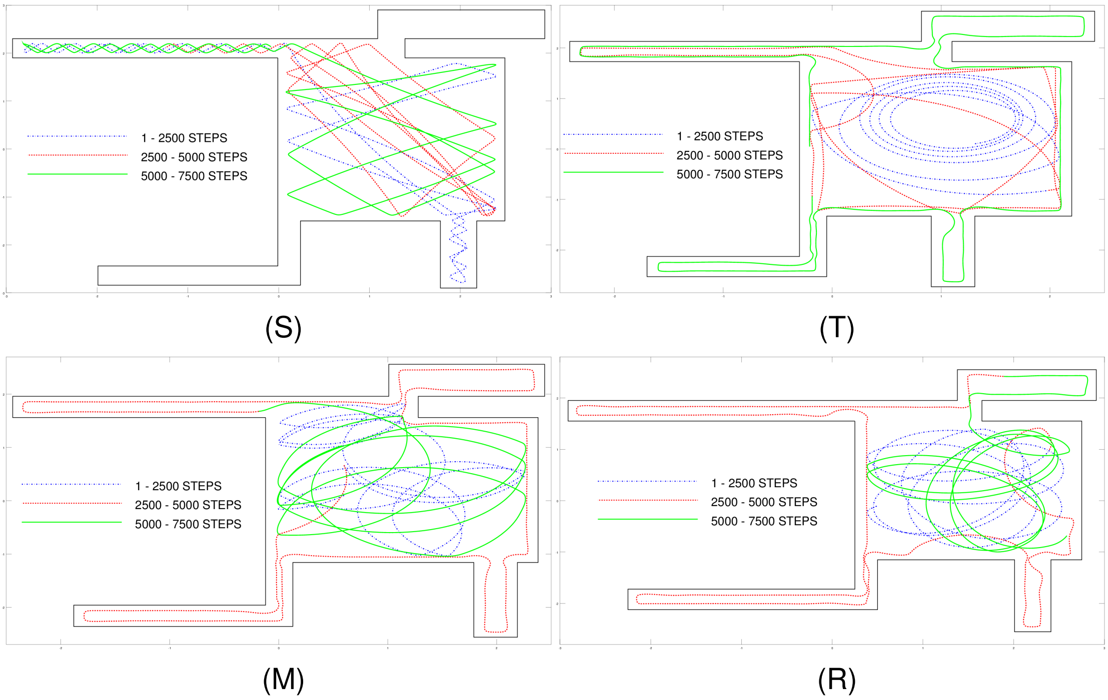


Figure 3.7: Typical trajectories displayed by the best robots of the four experimental conditions without hidden units in the concave environment. The portions of the trajectory produced during the first, second, and third part of the trial (i.e., from step 1 to 2500, from step 2501 to 5000, and from step 5001 to 7500, respectively) are shown with different colours and line style

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. 3.7, 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 carefully tune the time duration of the two behaviours. Indeed, the relative duration of the two behaviours determines whether or not the robot spends enough time exploring the central large area while keeping enough time to explore all the peripheral areas of the environment.

A qualitative analysis of the first ten replications showed that in the best two robots, that clearly statistically outperform the best robots of the other eight replications, the transition between the two behaviours occurs at 3.17±0.11min. 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 4.17min, when the robot operates on the basis of the first and third neural modules, and a wall-following behaviour displayed during 2.08min in which it operates on the basis of the second neural module (Fig. 3.7, 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 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. 3.7, 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 (2.99±1.02min). The best robots of four other replications displayed qualitatively similar solution, while the best robot of the five remaining replications displayed 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 achieved similar performance in all experimental conditions (see Fig. 3.8, left). The differences among the four experimental conditions are statistically 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. 3.8, 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. 3.9). 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.

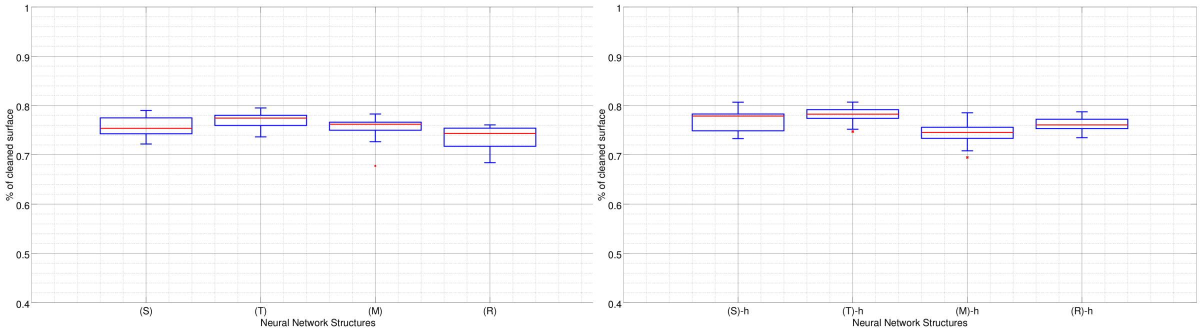


Figure 3.8: Box plots of performance in the convex environment. The left and right figures report the results obtained without internal neurons and with internal neurons, respectively

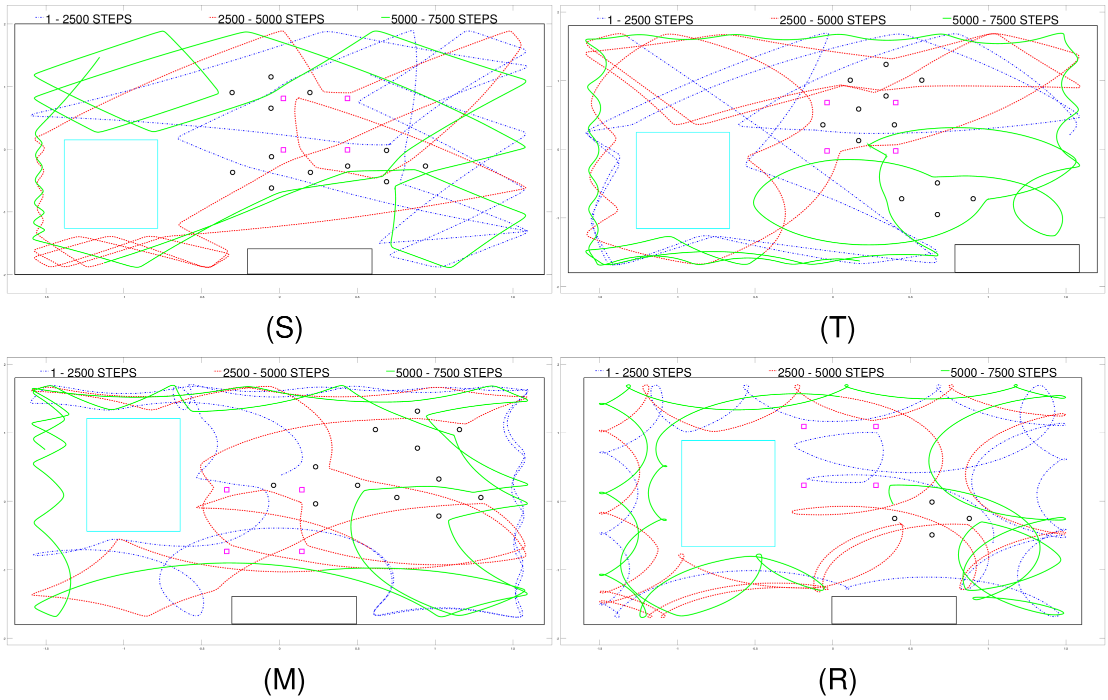


Figure 3.9: Typical trajectory displayed by the best robots of the four experimental conditions without hidden units in the convex environment

3.3.2 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 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 [55] 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 in 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. 3.10). This implies that, at this short timescale, all robots of all experimental conditions display behavioural plastic behaviours. The reasons that explain 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 keep exploring the environment) and it is supported by the availability of always-available and easy-to-use affordances. Indeed, independently of 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

Figure 3.10: 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

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 and the greater the range of timescales at which the robot exhibits differentiated behaviours, the greater the behavioural plasticity of the robot. In the remainder of this chapter, 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 longer 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 detail 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 [185, 188]. 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 [38], a task-independent learning process that can enable situated robots to synthesise temporarily stable behaviours, though the mechanism and the processes through which this is realised are completely different. Notice that the dynamical systems theory terms used here are referred to behavioral trajectories and not to the underlying neural mechanisms.

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. 3.7, 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 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. 3.12). 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. 3.13). 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 experience both 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. 3.11, 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 produc-

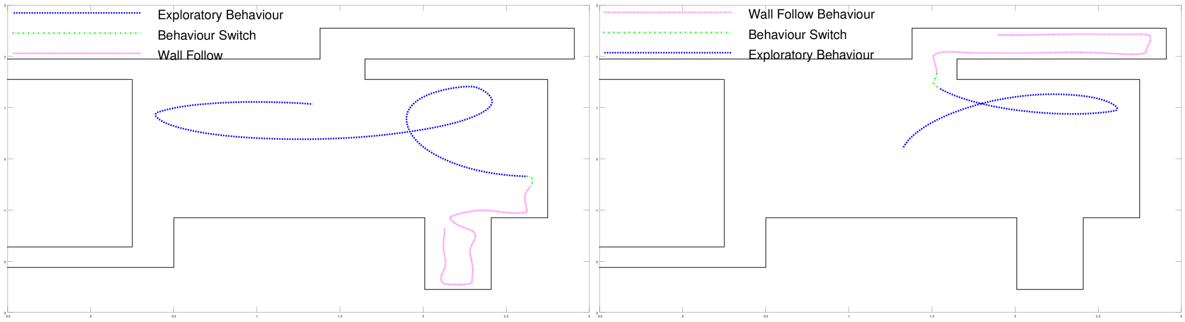


Figure 3.11: Illustration of how the best (R) robot switches from the exploration to the wall-following behaviour and vice versa (left and right, respectively)

tion 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±1.02min). The high variability negatively impacts its 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. 3.11, 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 2.08min, it is not surprising that behavioural switching occurs primarily during the switch between the

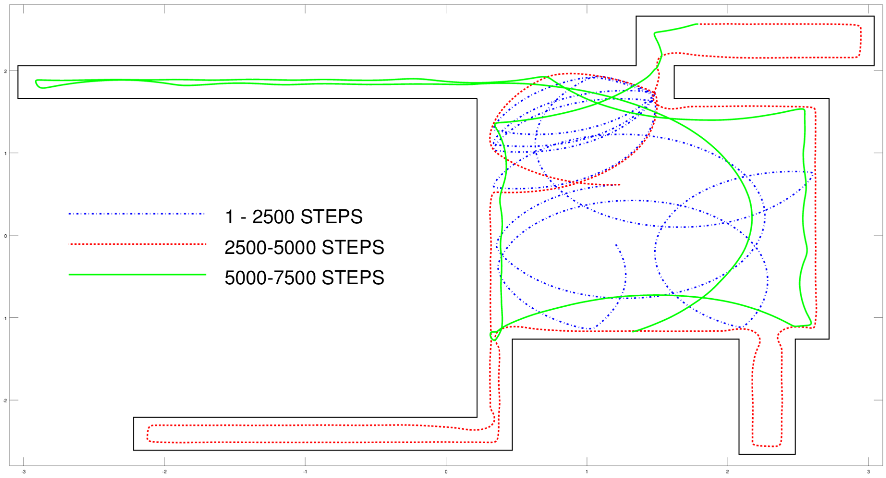


Figure 3.12: Trajectory produced by the best (M) robot which produces an exploration behaviour under the control of the first neural module, an exploration and then a wall-following behaviour under the control of the second neural module, and a wall-following and then an exploration behaviour under the control of the third neural module

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 4.17min and the wall-following behaviour for 2.08min 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. 3.12). 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 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±0.11min). The double regulation 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. 3.13). As shown in Fig. 3.13, whether or not the robot switches to the wall-following behaviour depends both on the state of the internal clock and the state of the infrared sensor that the robot experiences when it approaches the wall. Overall, this shows that whether or not the switch between the two behaviours occurs depends both on the state of the internal clock and the way in which the exploration behaviour is realised which, in turn, influences the type of stimuli experienced by the robot. As mentioned above, in the case of the best (T) robot, the state of the time neuron is used not only 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 also 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

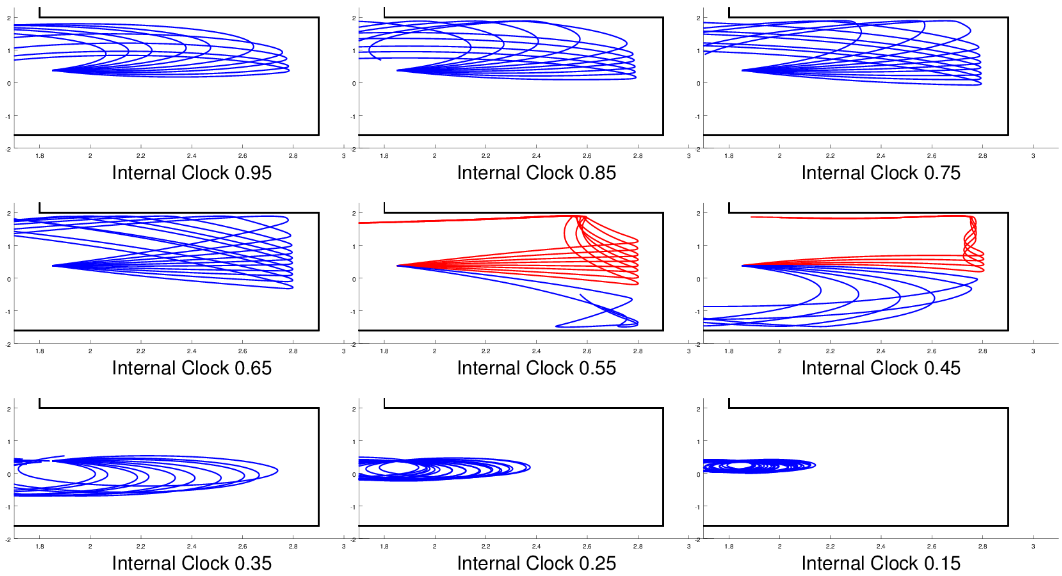


Figure 3.13: Behaviour produced by the best (T) robot during different trials in which it started from the same initial position with systematically varied orientations and systematically varied state of the time neuron. The red and blue lines represent the trajectories produced by the robot during trials 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

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. 3.14). 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.

Fig. 3.14 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

The smartest solution to the transition problem is that discovered by the best (T) robot (see Fig. 3.7, 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 from

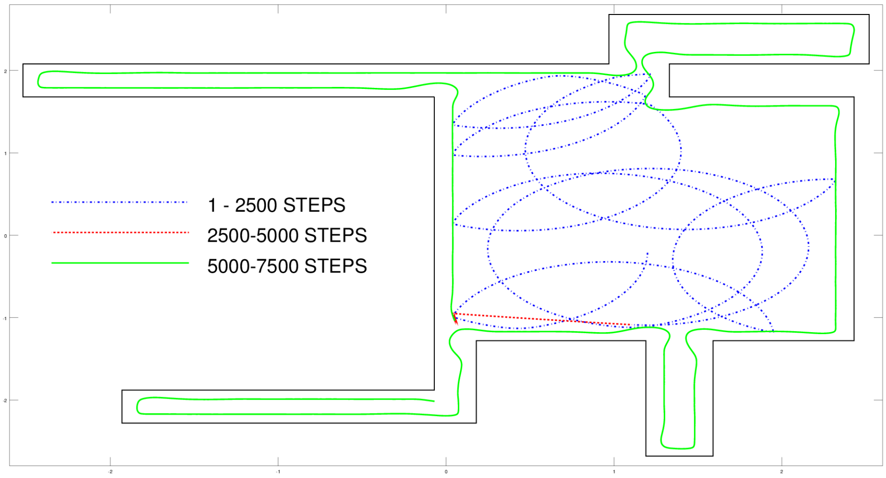


Figure 3.14: Trajectory produced by one of the three best (M) robots characterized by a second module that is specialized for enabling a suitable transition from the exploration to the wall-following behaviour.

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

3.4 Discussion

In this chapter a detailed study showing how behavioural plasticity can emerge and can make the evolving robots more adaptive was presented. Furthermore it was demonstrated that it can emerge independently from whether or not the task requires to face multiple conflicting goals, and also without the necessity of structural changes or additional selective pressures that direct the process to evolve any kind of modularity. 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.

Interestingly, the behaviours displayed by the best evolving robots show similarities with those obtained in [61] 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 [105, 130, 150]), consists in facilitating the synthesis of behavioural plastic solutions.

The analysis of the obtained results indicates that the mechanisms supporting 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 among different behaviours.

The perception of affordance constitutes a prerequisite for the possibility to develop differentiated behaviours and effectively arbitrate them, i.e., selecting the appropriate behaviour for the current robot/environmental context and regulating the duration of each behaviour. Interestingly, the basic mechanism 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. While the importance of affordance perception and usage is widely recognised, the notion of affordance generation introduced in this study and the description of how affordance generation supports the evolution of multiple context-dependent behaviour are original contributions of this work.

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. 3.7, 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. 3.13).

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 from both an evolutionary perspective, since it enables a gradual transformation [34, 84], and 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 the presented experiment).

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 [185], 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 complex behaviours is often 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 [7, 23, 71, 166]. 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 [66, 169]. From this perspective, the results presented in this chapter 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.

Chapter 4

Cognitive Offloading Does Not Prevent But Rather Promotes Cognitive Development

In this chapter a study investigating the factors that can promote the evolution of integrated cognitive behaviours in evolving robots are analised. Furthermore, the investigation helps to clarifity how reactive strategies, those relying solely on immediately perceived stimuli, can be combined with cognitive strategies, those relying on internal information, for creating a hybrid and robust strategy that leads to the task solution.

The experimental setup chosen included a delayed response task, the double t-maze, and robots with different neural network architectures. As we will see, the results obtained indicate that the use of a reactive strategy called cognitive offloading, namely the development of an ability to encode information extracted from the sensory-motor states into the external environment, does not prevent but rather promote the evolution of truly cognitive abilities. More generally we will show how the development of behavioural capabilities that rely on direct sensory-motor rules constitute a prerequisite for the development of cognitive capabilities that rely also on internal states.

4.1 Introduction

Developments in psychology, neuroscience, linguistics, robotics and philosophy have clarified that cognition cannot be studied properly without taking into sufficient account the role of the body, action and the external world [30, 31, 135, 137, 165, 170, 177]. The agent's body and the environment in which it is situated provide a great deal of structure that is used to operate appropriately. Consequently, in many cases the internal capabilities required are much simpler than those previously hypothesized within disembodied accounts. For example, moving around in a city does not necessarily require an elaborate representation of the city's layout. The ability to recognise a limited number of turning decision points combined with the ability to just follow the street between decision points might suffice [83]. Similarly, baseball players do not need to estimate the trajectory of the flying ball to be intercepted through complex calculations. They can simply adjust their running speed so as to maintain the relative angle between their eyes and the ball constant [108].

Actually embodied cognition is an alternative approach for thinking about cognition that, instead of considering it as a high level, completely internal, process operating over abstract symbols, considers it as a situated activity meaning that the external actions are part of the cognitive process [2]. The above examples are just some cases that are much better explained by the embodied cognition theory than by the traditional computational theory of mind [72]. Some authors go beyond and claim that even external technological artifacts can be considered as an extension of the mind [32]. This embodied way of looking at cognition creates many new unexplored possibilities, considering that actions can exploit the regularities present in the environment, and in the sensory-motor loop. If action is indeed part of cognition it would be important to better understand its role in the cognitive process.

The present study aims to investigate the relation between the development of reactive and cognitive capabilities. In particular, the objective is to investigate whether the development of reactive capabilities prevents or promotes the development of cognitive capabilities. For the scope of this study cognition is defined as the ability to integrate sensory-motor information over time into internal states and to use these internal states to regulate the way the agent reacts to perceived stimuli. The above definition is focused on a fundamental capacity that is at the basis of all cognitive capabilities (e.g. perception, memory, attention, decision-making, reasoning, language, etc.).

Evolutionary Robotics [120] is a suitable method for studying the relation between reactive and cognitive abilities in adaptive agents. Indeed, research carried out in this area has demonstrated how evolving robots can master both problems that have reactive solutions and problems that require the development of cognitive abilities (see for example [13, 14, 29, 48, 57, 62, 65, 155, 195]). However, what has not been sufficiently investigated to date is the relation between reactive and cognitive strategies.

Adaptive problems typically admit qualitatively different sub-optimal and optimal solutions. Depending on the circumstances, the discovery of one type of sub-optimal solution might facilitate or block the discovery of better alternative solutions. Indeed, the discovery of sub-optimal solutions that cannot be progressively transformed into better solutions without loss of performance should retard or block the discovery of effective solutions. On the contrary, the selection of sub-optimal solutions that can be transformed into better solutions without causing significant performance loss can facilitate the discovery of effective solutions. In the latter case, the strength of the facilitation effect would depend on the level of overlap or similarity between the first and the second solutions.

Since reactive solutions are typically simpler than cognitive solutions from the point of view of the complexity of the required control mechanisms, and since adaptation tends to find the simpler solutions first, the question is the following: Can reactive solutions be transformed into better solutions that also include the utilisation of internal states without causing significant performance loss and what is the level of overlap/similarity between sub-optimal reactive solutions and better solutions that include cognitive capabilities?

According to some authors, the development of reactive solutions blocks the development of cognitive capabilities. In particular, [127] claimed that reactive solutions constitute hard to escape local optima that prevent the development of cognitive solutions. Similarly, [95] claimed that the deception caused by the availability of locally optimal reactive policies is one of the main factors that explains why it is difficult for cognitive policies to evolve. For these reasons the authors hypothesised that the development of cognitive solutions necessarily requires specific selective cognitive pressures such as fitness components that encourage the development of short-term memory [127] or mechanisms for avoiding local optima, such as novelty search [95].

One aspect that is particularly relevant from the viewpoint of the relation between reactive and cognitive strategies is cognitive offloading, i.e. the possibility of offloading cognitive work onto the environment [101, 144, 190]. More specifically, the possibility of acting so as to encode the states that can be used to regulate the agent's behaviour onto the external environment and/or onto the relation between the agent and the environment. Indeed, the possibility of encoding the required states internally or externally suggests that cognitive strategies and reactive strategies (that rely on cognitive offloading) represent two alternatives but functionally equivalent modalities. A simple example of cognitive offloading related to everyday human life is crossing two fingers so to avoid forgetting to perform a certain action [58, 59, 144]. An example of cognitive offloading realized in a robotic scenario consists on dropping markers in the environment in order to use them to find the way back to the home location [29].

Cognitive offloading is usually considered in the case of cognitive agents that already possess cognitive abilities but offload information into the environment or into their relation with the environment. In this work, instead, we focus on a perspective in which the agents need to evolve a certain skill to adapt to their environment and can do so by using a reactive strategy that relies on cognitive offloading, a cognitive strategy that relies on internal states, or on a hybrid strategy that relies on both.

In the context of this work a series of experiments was performed in which a population of robots provided with neural network controllers were evolved for the ability to master a navigation problem in a double T-maze environment that required the exhibition of delayed response behaviour. Analysis of the experiments revealed that the evolving robots always selected reactive strategies that relied on cognitive offloading. The discovery of these strategies did not prevent but rather facilitated the development of improved strategies that also relied on the extraction and use of internal states. A detailed

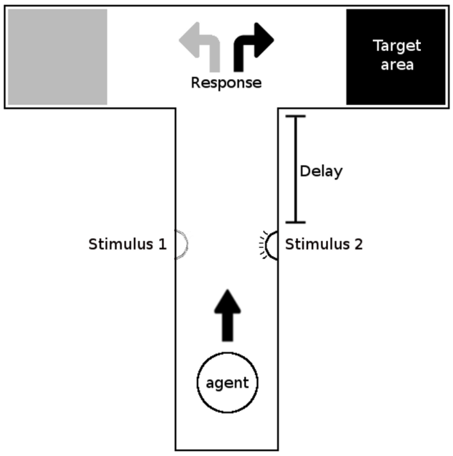


Figure 4.1: The T-Maze task. The bottom portion of the central corridor includes a stimulus located on the left or the right side. The position of the stimulus corresponds to the position of the target destination. Adapted from [99].

analysis of the results obtained in different experimental conditions provides evidence that helps to clarify why, contrary to expectations, reactive and cognitive strategies tend to have synergetic relationships.

4.2 Method

A paradigmatic class of problems that require cognitive skills is constituted by delayed response tasks in which an agent has to act at a certain time t in a conditional dependent manner on the basis of stimuli it encountered at a previous time (t - delay). A simple example of a delayed response task is the so-called road sign task in which an agent that is initially located at the bottom of a T-Maze environment needs to travel toward the top-left or top-right destination by turning left or right at the T-junction depending on whether it previously experienced a stimulus on the left or on the right side of the central corridor, respectively (Figure 4.1). Therefore, this task was chosen by several researchers to study the evolution of cognitive robots [4, 9, 41, 99, 127, 147].

Actually, as demonstrated by Ziemke and Thieme [195] this problem has a non-cognitive solution, i.e. a reactive strategy in which the robot always acts on the basis of its current sensory state. Indeed, the robot could solve the task by approaching the experienced stimulus, when visible, and then by following the nearby left or right wall. The availability of simple and optimal reactive solutions of this type prevents the development of cognitive solutions. These reactive solutions, indeed, are easy to discover, optimal and, consequently, evolutionarily stable.

The question that remains open is whether cognitive solutions can evolve in experimental settings that do not allow for optimal reactive solutions or whether the discovery of sub-optimal reactive solutions prevents the discovery of better solutions [95, 127] and consequently might drive the evolutionary process toward inescapable local optima. To investigate this question we decided to carry out the evolutionary experiments described in the following sub-sections.

4.2.1 The task, the environment and the robot

The environment consisted in a double T-Maze (Figure 4.2) that included four different destinations and two types of stimuli that could be experienced in four different corresponding patterns (left-left, left-right, right-left, right-right). The increase in complexity with respect to the simple T-Maze environment was due to the fact that the number of destinations to be reached was higher, the number of stimuli experienced was higher, the number of stimuli-dependent decisions that had to be made was higher and the time delay between the moment in which the stimuli were experienced and the moment in which the stimuli-dependent decisions had to be made was longer.

To obtain solutions that were robust with respect to environmental variations, the initial positions and orientations of the robot and the size of the environment were randomly varied at the beginning of each trial. More precisely the length of the central corridor and the two vertical corridors was randomly set during each trial within 4.5m±0.5m and 5.5m±0.5m, respectively. The position between the two signals was also varied proportionally with the length of the central corridor. The initial position of the robot was selected randomly within a 50x45cm rectangular area located at the beginning of the central corridor. The initial orientation of the robot was selected randomly with a uniform distribution in the [-180, 180]° range. Robots were evaluated for 16 or 32 trials,

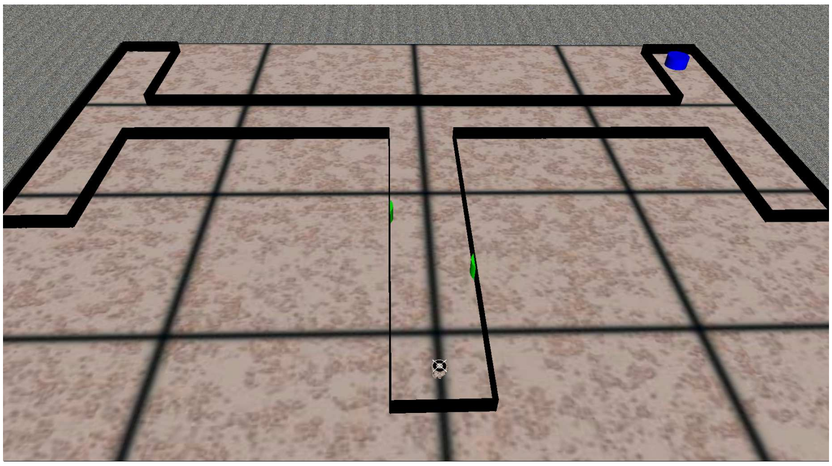


Figure 4.2: The double T-Maze task. The blue cylinder, which represented the target destination, was located at one of the four end points of the lateral corridors. The central corridor included two green stimuli located in the first and the second part of the corridor on the left or the right side. The position of the first and the second stimulus indicated whether the robot should turn left or right at the first and the second junction, respectively, to reach the target destination.

as explained below. Trials lasted up to 1 minute and were stopped as soon as the robot turned in the wrong direction at one of the two junctions.

Complex T-Mazes have already been used in previous research. In particular [19] evolved robots for the ability to navigate in T-Maze environments in multiple trials in which the destination location remained stable. The authors manage to develop robots that were able to “remember” the target location in simple T-Mazes but not in double T-Maze environments. [131] evolved robots for the ability to navigate in a triple T-Maze toward 8 alternative destinations. In these experiments, however, the robots were not required to master a delayed response task. Indeed, they received and had access to one of eight different corresponding patterns for the entire duration of each trial.

We used the MarXbot [20] agent, which is a circular robot with a diameter of 17cm equipped with 24 infrared sensors, a rotating scanner sensor and an omnidirectional camera. The experiments were run in simulation by using the FARSA open-software tool that includes an accurate simulator of the robot and of the environment [106, 107]. FARSA has been used to successfully transfer results obtained in simulation to hardware in similar experimental settings (e.g., [62, 156]).

4.2.2 The robots' neural controllers

Evolving robots were provided with neural network controllers. The sensory layer included eight sensory neurons that encode the average activation state of eight groups of three adjacent infrared sensors, 6 neurons that encoded the average activation of the rotating scanner over 60 degrees, and 12 neurons that encoded the percentage of red, green and blue pixels detected in four 90° sectors of the visual field of the camera. The state of the sensory and motor neurons was normalized in the range [0.0, 1.0] and noise is simulated by the addition of random values selected with a uniform distribution in the range [-0.05, 0.05]. The motor layer includes two motor neurons that encode the desired speed of the two corresponding motors (normalised in the range [-10.0,10.0]cm/s) that actuate the differential driving system of the robot.

The experiments were replicated in three different experimental conditions that varied with respect to the architecture of the robots' neural controller as described below:

(S) Simple: The robots were provided with a simple reactive neural network (that always responded in the same way to the same sensory state) in which the sensory neurons were directly connected to the motor neurons (see Figure 4.3).

(C) Continuous: As in the case of the previous condition the neural network controller included direct sensory-motor connections. In addition, the network included an internal layer with 6 neurons that received connections from the sensory neurons and projected connections to the motor neurons. The internal neurons were continuous, i.e. their output state depended on both the activation received from the incoming connections and on their previous activation state (see [15, 56]) (see Figure 4.4).

(CR) Continuous Recurrent: The neural network was identical to the previous condition, but the internal neurons were also interconnected through recurrent connections (see Figure 4.5).

The state of the sensors, the neurons and the desired speed of the motors were up-

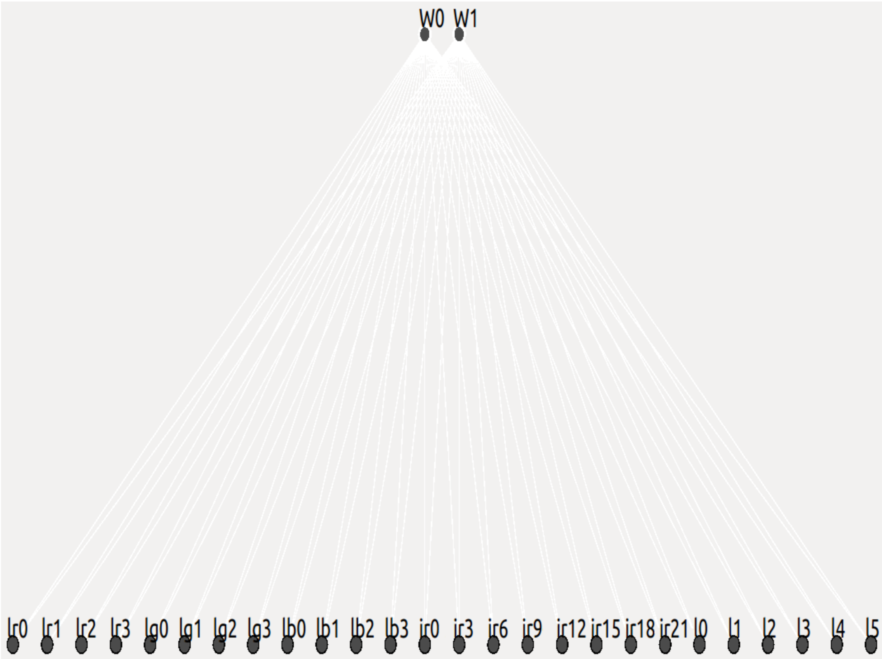


Figure 4.3: Simple Neural Network Controller (S). The output neurons W define the motor activation for left and right wheels of the robot. Neurons with a white contour around it mean the presence of a bias parameter.

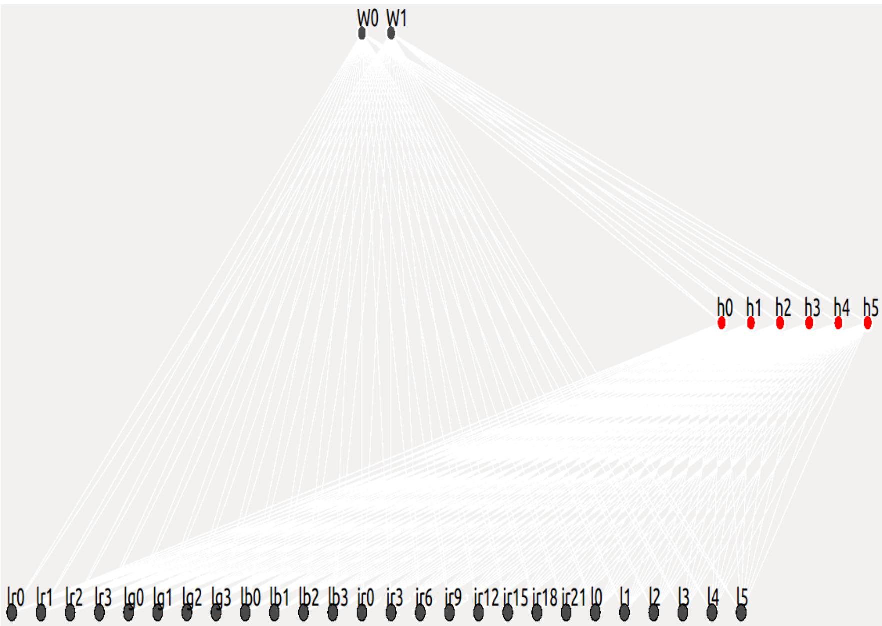


Figure 4.4: Continuous Neural Network Controller (C). Red filled neurons mean continuous neurons. Neurons with a white contour around it mean the presence of a bias parameter. The output neurons W define the motor activation for left and right wheels of the robot.

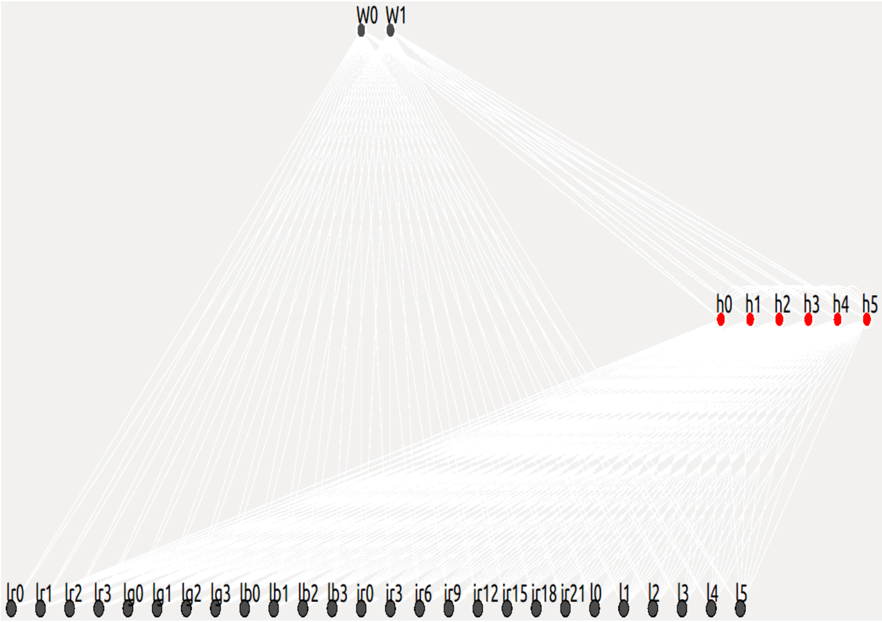


Figure 4.5: Continuous Recurrent Neural Network Controller (CR). Red filled neurons mean continuous neurons. Neurons with a white contour around it mean the presence of a bias parameter. The output neurons W define the motor activation for left and right wheels of the robot.

dated every 50ms. The architecture of the neural network controller was kept fixed. These experimental conditions were chosen to enable us to verify whether and to what extent the problem had reactive solutions, whether the possibility of integrating sensory-motor information over time into internal neuron states would enable the robots to develop better solutions and whether reactive strategies could coexist with cognitive strategies.

4.2.3 The evolutionary algorithm

The initial population consisted of 20 randomly generated genotypes that encoded the connection weights, biases, and time constants of the neural network controllers of 20 corresponding individual robots. Each parameter was encoded with 8 bits and normalized in the range [–5.0, +5.0] in the case of connection weights and biases and [0.0, 1.0] in the case of time constants of continuous neurons.

Each individual was allowed to generate one offspring, i.e. a copy of the parent genotype in which each bit was mutated with a given probability. Each offspring was evaluated and was used to replace the genotype of the worst parent or was discarded depending on whether or not it outperformed the worst parent.

Each experiment was repeated 10 times by starting with different randomly initialised genotypes. The evolutionary process continued in two consecutive phases of 2,000 generations. During the first 1,000 generations, the mutation rate was set at 2% and the evolving robots were evaluated in 16 trials. From generation 1,001 on the mutation rate was set at 1%. From generation 1501 to 2000 on the individuals were evaluated in 32 trials. In some of the experiments the robots were subjected to position and orientation noise during the second phase, as described below.

The fitness of the individuals was computed by averaging the fitness obtained during the different trials. The fitness of each trial corresponded to the inverse of the distance, within the maze, between the robot and the target destination at the end of the trial. In other words, the robots were rewarded for the ability to approach the appropriate destinations. Equations 4.1 and 4.2 present the fitness equation used where T is the total number of trials in each generation, which as presented above can be 16 or 32, d_i is the distance within the maze between the robot and the target at the beginning of the trial, and d_f is the distance within the maze between the robot and the target at the end of the trial.

$$fitness = \sum_{t=1}^{T} \frac{1}{T} * f_1 \tag{4.1}$$

$$f1 = \begin{cases} 0 & \text{if } d_f > d_i, \\ \frac{d_i - d_f}{d_i} & otherwise \end{cases} \tag{4.2}$$

The experiments described in this chapter can be replicated by downloading and installing FARSA and the associated experimental plugin from
“https://sourceforge.net/projects/farsa/” and from
“http://laral.istc.cnr.it/cognoffpone/dtmaze.tgz”.

4.3 Results

In this section we report the results obtained in the different experimental conditions described in Section 4.2 and in additional control experiments described below that were carried out to clarify the role of cognitive offloading in the development of cognitive skills.

Overall the performance of the robots (i.e. the percentage of trials in which the evolved robots reached the correct target destination) did not differ significantly between the three experimental conditions in which they were provided with different neural controllers (see Figure 4.6). By analysing the performance of the best robot obtained in each experimental condition (Figure 4.6) we can see how the best (C) and (CR) robots managed to achieve close to optimal performance in 92.8% and 96.3% of the trials, respectively, while the best (S) robot only achieved sub-optimal performance (i.e. 71.7% success). The performance of the best (C) and (CR) robots did not differ significantly (Fisher Exact Test, $p=0.308$). The performance of the (S) robot, instead, was significantly worse than the performance of the (C) and (CR) robots (Fisher Exact Test, $p<0.001$). The data were obtained by post-evaluating the best robot from the last generation of each replication in 600 trials.

These results confirm that, at least in our experimental setup, the double T-Maze problem cannot be solved through simple reactive solutions, although the best reactive robot (S) achieved a remarkably high performance (i.e. by navigating toward the correct destination in 71.7% of cases). Moreover, these results indicate that the possibility of integrating sensory-motor information over time through continuous neurons (C) and recurrent connections (CR) into internal states that are used to regulate the way the robots react to sensory stimuli enables the evolving robots to achieve close-to-optimal performances. On the other hand, the fact that close to optimal performances were achieved in a minority of the replications indicates that the task is hard and there is a high probability that evolution remains stuck in sub-optimal regions of the search space.

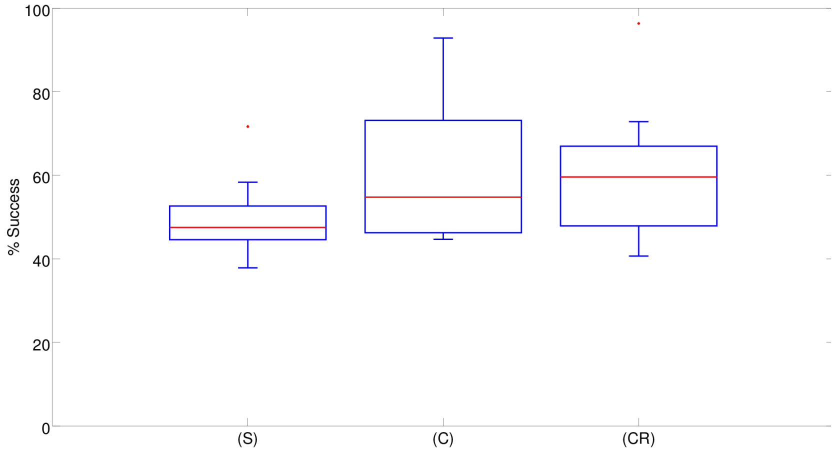


Figure 4.6: Performance (i.e. percentage of successful trials) of the best 10 robots evolved in the (S), (C) and (CR) experimental conditions in 10 corresponding replications of each experiment. Boxes represent the inter-quartile range of the data and horizontal lines inside the boxes mark the median values. The whiskers extend to the most extreme data points within 1.5 times the inter-quartile range from the box. Circles indicate the outliers.

By inspecting the trajectories produced by the best (S) robot, i.e. by means of a simple reactive controller, we can see how the robot navigated correctly toward all four target destinations most of the time, but erroneously navigated toward the left-left and right-left destinations instead of the left-right destination in several trials (Figure 4.7, blue trajectories). Surprisingly, this shows that the double T-Maze task can also be solved to a large extent with a reactive solution in which the robot's actions depend only on the current robot's input and in which the robot does not store any internal information regarding previously experienced sensory states. This is achieved by offloading the critical information into the environment (or more precisely into the robot/environmental relationship).

Indeed, the behavioural analysis of the best (S) robot indicates that the experienced signals are used to systematically alter the positions assumed by the robot at the end of the central corridor (Figure 4.7). These positions influence the type of stimuli the robot experiences at the first junction which, in turn, determine whether the robot will turn left or right at the junction. The position of the robot at the end of the first corridor also influences how the turning is realised, i.e. whether the robot produces a tight turn

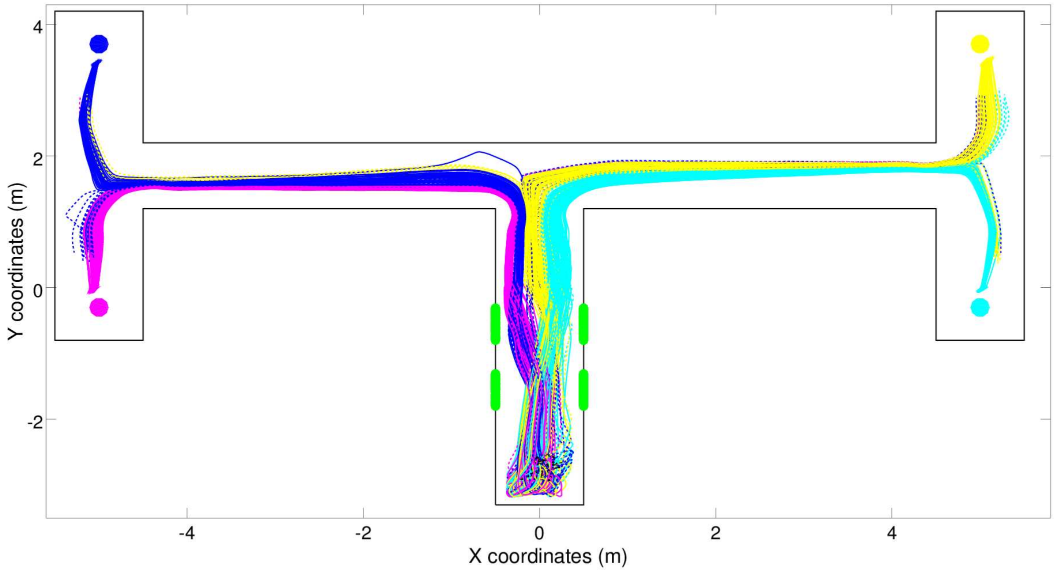


Figure 4.7: Trajectories produced by the best (S) robot over 300 trials. Full and dashed lines indicate successful and unsuccessful trials, respectively. The color indicates the corresponding target destination (magenta: left-bottom, blue: left-top, cyan: right-bottom, and yellow: right-top).

or a wider one, and consequently the position assumed by the robot in the second corridor. Indeed, after experiencing the right-right signals the robot assumes the right most position at the end of the first corridor and then a right position in the second corridor. By contrast, after experiencing the right-left signals the robot assumes the central position at the end of the first corridor and then a left position in the second corridor. This ability to differentiate the relative position assumed in the second corridor on the basis of the position assumed at the end of the first corridor enables the robots to turn in the appropriate direction also at the second junction. The same things happen when the robots travel towards the other two left destinations.

This interpretation is confirmed by the result of an analysis in which the robot was initially placed at the end of the first corridor with an orientation that systematically varied in the range of [-55, 55]° with respect to the orientation of the first corridor and a position that systematically varied along the x-axis between [-0.4, 0.4]m with respect to the center of the corridor. As shown in Figure 4.8, the destination reached by the robot depended primarily on the relative position along the x-axis and secondarily

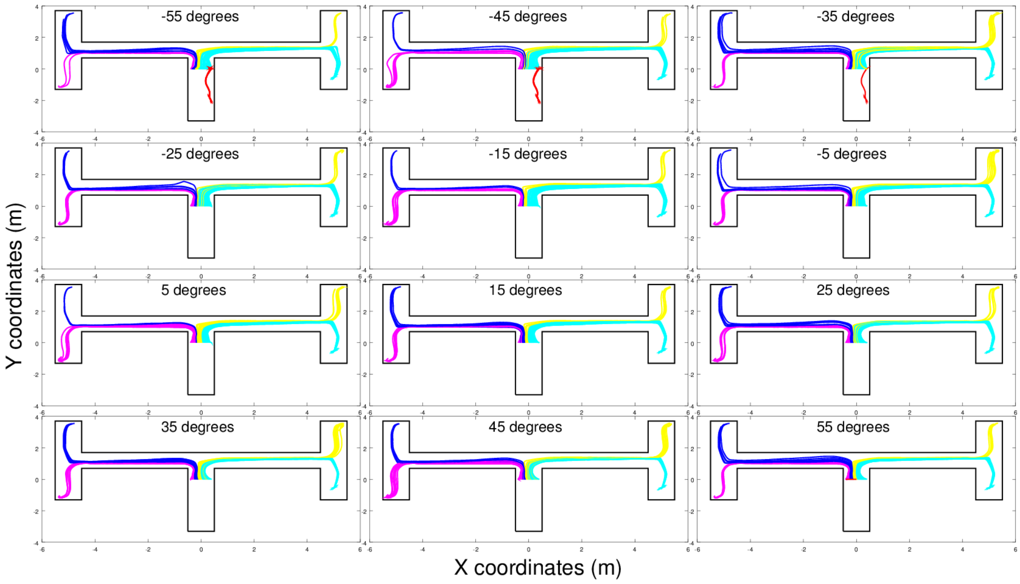


Figure 4.8: Behavior displayed by the best (S) robot placed initially at the end of the first corridor in different positions (along the x-axis) and orientations. The color of the trajectories corresponds to the destination reached (magenta: left-bottom, blue: left-top, yellow: right-top, and cyan: right-bottom). Trajectories that did not enable the robot to reach any target destination are shown in red.

on the orientation of the robot at the end of the first corridor. This means that the robot offloads the information encoding the destination to be reached in its position and orientation. The ineffective behaviours shown in red, which are produced when the robot is initially located near the right wall and oriented toward the right, occur only occasionally in normal conditions (Figure 4.7) because the robot rarely reaches these positions/orientations when it starts from the beginning of the central corridor.

The fact that the destination reached by the robot depended on the position and the orientation of the robot at end of the first corridor is demonstrated by the fact that the destination the robot will reach at the end of the trial can be predicted with 84% accuracy on the basis of the position and the orientation the robot assumes in the first corridor, 40cm before the first junction. This success rate was obtained by training a feed-forward neural network through a backpropagation algorithm based on a cross-entropy error function [114]. The network included two inputs neurons that encoded the x-position and the orientation of the robot, six hidden neurons and four winner-take-all outputs neurons that encoded the four corresponding destinations. The training set consisted of 6,000 position and orientation vectors and 6,000 corresponding destination vectors. The fact that the target destination could not be predicted in all cases can be explained by considering the effects of noise on sensors and motors. Further evidence demonstrating that the target destination reached by the best (S) and (C) robots depended on the position and the orientation assumed by the robots from the end of the first corridor on are reported below.

The four behaviours displayed by this robot (indicated by the trajectories shown in magenta, blue, cyan and yellow in Figure 4.7) are dynamical processes that arise from the robot/environmental interactions and that converge toward four fixed-point attractors. This can be appreciated by observing the four corresponding basins of attraction in the 2D projection of the phase portrait (Figure 4.9). These basins of attraction enable the robot to reach four different destinations without varying the way it responds to perceived stimuli (i.e. by using a reactive controller that always responds in the same way to the same stimuli independently of the stimuli experienced before). This can be explained by considering that the way in which the robot reacts to perceptual stimuli and the way in which perceptual stimuli change (as a function of the action performed by the robot and the characteristics of the local portion of the environment) ensure that the robot keeps moving towards the current destination while remaining in the current basin of attraction. To solve the problem, therefore, the robot only needs to enter into the appropriate basin of attraction in the first corridor while it perceives the green stimuli. Notice that the dynamical systems theory terms used here and in the rest of this chapter are referred to behavioral trajectories and not to the underlying neural mechanisms.

The basins of attraction also ensure that the behaviour of the robot is robust with respect to perturbations caused by noise and environmental variations (within limits). This is because typically small alterations in the robot's position and orientation do not

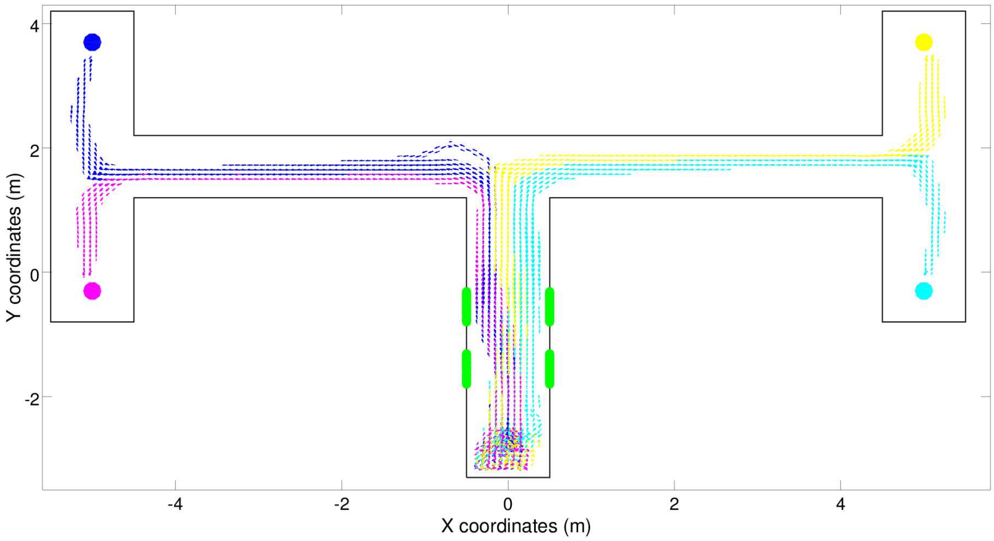
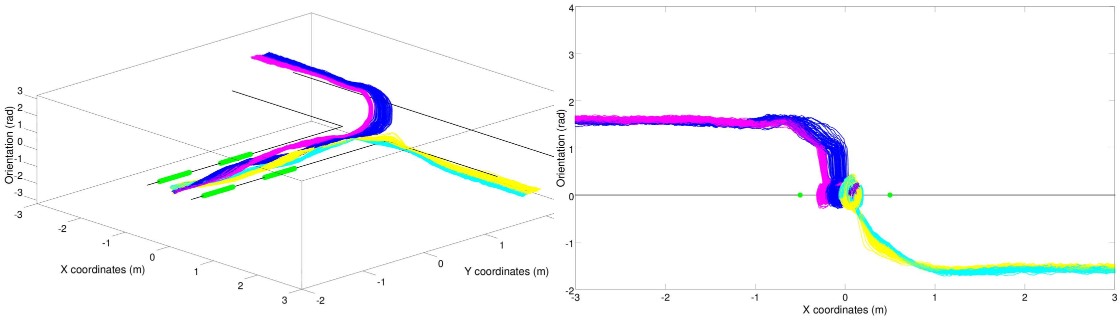


Figure 4.9: 2D vector field displaying the velocity of the robot in varying positions of the x/y plane for the (S) robot. For the sake of clarity, the vectors are shown only for the positions and orientations reached by the robot in natural circumstances (i.e. the positions and orientation assumed by the robot in 300 trials). The multiple arrows displayed in each 75x75mm position cell indicate how the direction and the magnitude of the velocity vector varies as a function of the different orientations assumed by the robot in the corresponding position.

cause a switch from one basin of attraction to another and consequently do not alter the robot's destination. Moreover, this is also because the effects of small alterations tend to be automatically compensated over time by the convergent nature of the attractors that drive the robot away from the borders that separate the different basins of attraction. Note that, as shown in Figure 4.10, the state space includes three dimensions (x position, y position and orientation). Consequently, the four basins of attraction are separated also in areas in which they seem to overlap from the perspective of the 2D projection shown in Figure 4.9.

Selection of the appropriate behaviour (i.e. the convergence toward the appropriate basin of attraction) is the result of the bifurcation process that occurs in the first corridor and that is regulated by perception of the green stimuli. In other words, it is the result of the fact that, while the robot travels along the first corridor, it varies its position and orientation on the basis of the perceived green stimuli in a way that ensures that at the end of the first corridor the robot assumes a position and orientation that enable it to



(a) (b)

Figure 4.10: Phase portrait of the robot/environmental dynamics in the case of the best (CR) robot. Data collected during in 300 trials. The vertical axis represents the orientation of the robot in the range [-90, 90]° with respect to the direction of the first corridor. For the sake of clarity the plots refer only to the first-junction portion of the environment. (a) displays a view from which variation on the three dimensions can be appreciated. Instead, (b) shows a 2D orthogonal projection in which one can appreciate only variations along the vertical and horizontal x-axis that correspond to the orientation of the robot and to the position of the robot along the x-axis, respectively.

enter in the right basin of attraction.

For example, let us consider the trials in which the robots experience the right-right signal (Figure 4.7, cyan trajectories). In these cases, the robot reacts to the two green stimuli located on the right by moving toward the left side of the corridor. This enables the robot to turn right at the first junction, because it turns right when it encounters a wall ahead and a wall on its right side that is nearer than the wall on its left side, and then to assume a specific position and orientation in the second corridor that enable it to turn right also at the second junction. Thus, the proper movements produced in response to the stimuli perceived in the first corridor ensure that the robot environmental/dynamics will enter into the basin of attraction of the right-right behaviour, that then guide the robot toward the appropriate destination. During the trials in which the robot experiences the right-left signal, instead, the robot reacts to the signals by moving first right and then left so as to assume a central position within the first corridor (Figure 4.7, yellow trajectories). This makes the robot turn right at the first junction also in this case. However, the right turn initiated from this position and orientation drives the robot toward the left side of the second corridor. This in turn enables the robot to then turn left at the second junction. In other words, the position/orientation with which the robot approaches the first junction influences not only whether it turns right or left, but also the position/orientation taken after the turn in the second corridor, which finally determines whether the robot will turn left or right at the second junction.

There are two reasons for the errors produced by the robot (Figure 4.7, dashed lines). The first is that in some cases the bifurcation process fails, i.e. the robot is unable to react to the stimuli experienced and, thus, assume the appropriate positions/orientations at the end of the first corridor. Consequently, the robot/environmental system enters into the wrong basin of attraction. This problem particularly affects some of the trials in which the robot experiences the left-right signals. Indeed, in 10.7% of these trials the best (S) robot erroneously navigates toward the right-left destination (Figure 4.7, top, dashed lines). This problem occurs when the robot starts from certain specific positions and orientations and/or as a result of noise or when the robot occasionally exits from the right basin of attraction and enters into another, wrong basin of attraction. This typically occurs as a result of noise in areas in which the divergence between two nearby basins of attraction is weak. In the case of the best (S) robot, this second type of problem occurs particularly in the left corridor. Indeed, in this phase the robots traveling toward the left-right destination erroneously enter into the behavioural attractor of the left-left destination in 46.7% of the left-right trials (Figure 4.7, dashed lines).

The behaviours displayed by the best (C) and (CR) robots are qualitatively similar to those displayed by the best (S) robot, see Figure 4.11. Indeed, also the robot/environmental dynamic of these robots is characterised by four fixed-point attractors (Figure 4.12). Moreover, the trajectories of these robots also bifurcate in the first corridor to ensure that the robot/environmental dynamic enters into the appropriate basin of attraction.

However, these robots make far fewer errors during the bifurcation phase than the (S) robot thanks to their ability to converge toward similar positions and orientations while they move along the first part of the first corridor (see Figure 4.11). Indeed, the

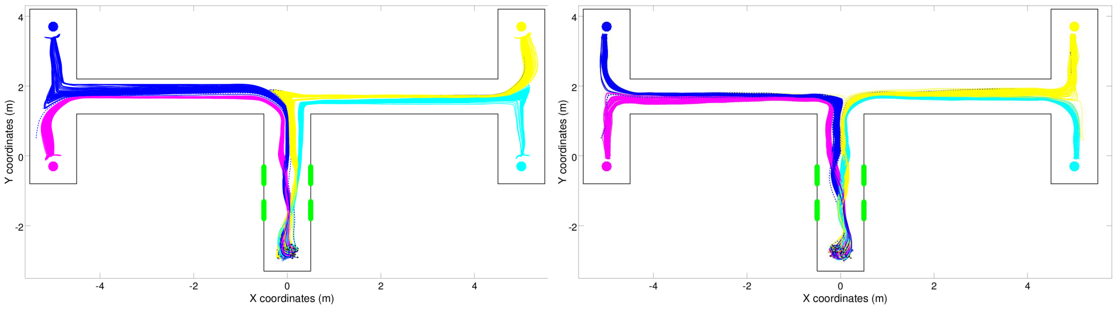


Figure 4.11: Trajectories produced by the best (C) and (CR) robots in 300 trials, left and right pictures respectively. Full and dashed lines indicate successful and unsuccessful trials, respectively. The color indicates the corresponding target destination (magenta: left-bottom; blue: left-top, cyan: right-bottom, and yellow: right-top).

variability along the x-axis of the positions assumed by (C) and (CR) robots one meter before the green stimuli is significantly lower than the variability of positions assumed by (S) robots (F-Test $F(299,299) = 3.114$, $p<0.001$ and $F(299,299) = 2.845$, $p<0.001$, respectively). By assuming a relatively un-variant position and orientation before they perceive the green stimuli, (C) and (CR) robots manage to achieve a more reliable bifurcation process than (S) robots.

Also, the errors that occur when the robots exit from their current basin of attraction and enter into another basin of attraction are reduced in (C) and (CR) robots. This is achieved through the synthesis of attractor basins, which produces a greater separation among the trajectories targeted toward different destinations that are produced by (C) and (CR) robots (Figure 4.11) than among the trajectories produced by the (S) robot (Figure 4.7).

Finally, in some cases the best (CR) robot also displayed the ability to re-enter into the basin of attraction in which it was previously located when it erroneously moved to another basin of attraction. This enables the robot to recover from some of the errors of this type caused by noise. This is demonstrated by the results collected during a post-evaluation test in which the robot was systematically displaced from its current basin of attraction to another one. The basin of attraction which the robot was displaced to

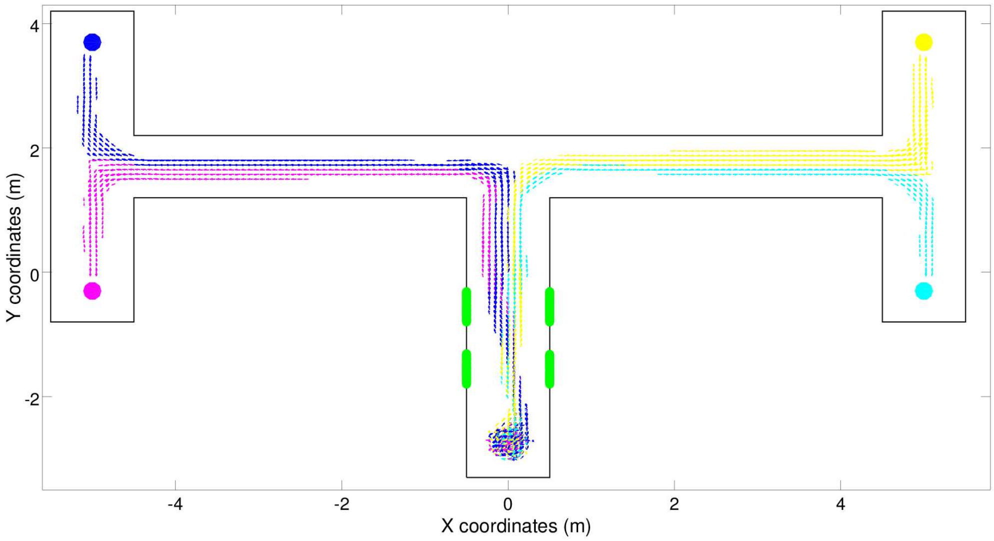


Figure 4.12: Vector field displaying the velocity of the robot in varying positions of the x/y plane for the (CR) robot. For the sake of clarity, the vectors are shown only for the positions and orientations reached by the robot in natural circumstances (i.e. the positions and orientations assumed by the robot in 300 trials). The multiple arrows displayed in each 75x75mm position cell indicate how the direction and the magnitude of the velocity vector varies as a function of the different orientations assumed by the robot in the corresponding cell and/or as a function of the robot's internal states.

was chosen randomly among the available ones, i.e. from the other three alternative basins of attraction before the first T-junction or between the only alternative attractor after the first T-Junction. The post-evaluation test was repeated in three conditions in which the robot was allowed to move normally and in which it was blocked for 1 or 3 seconds after the displacement. Analysis of the results indicates that the best (CR) robot was able to recover from this type of displacement in 60% of cases in which it was allowed to move normally after the displacement and in 20% of cases in which it was blocked in the displaced position and orientation for 1 second (Figure 4.13). The best (S) and (C) robots, instead, were able to recover from displacements only in a negligible percentage of cases (Figure 4.13). None of the robots were able to recover from displacements after being blocked in the displaced position and orientation for 3 seconds.

Overall the data reported in this section indicates that all robots offloaded information concerning the stimuli they experienced in the position and orientation they assumed

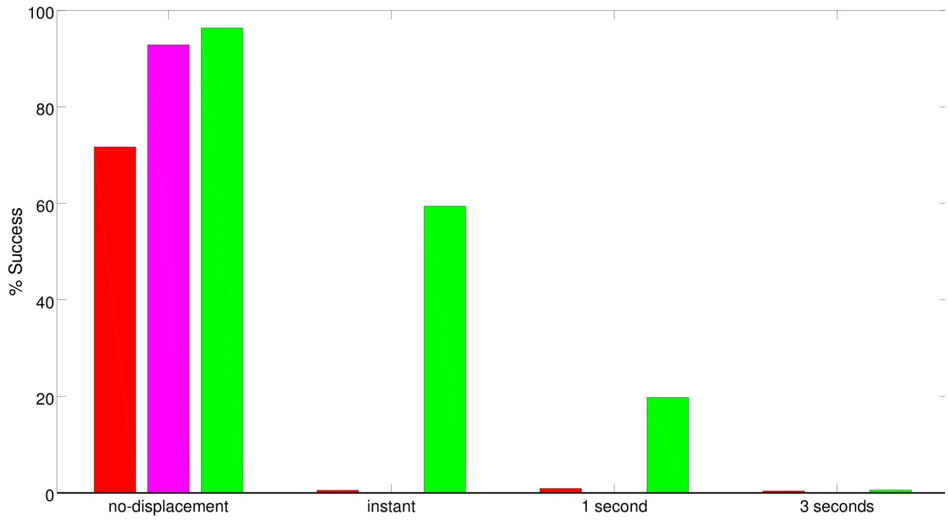


Figure 4.13: Performance displayed by the best (S), (C), and (CR) robots during a post-evaluation test in which they were randomly displaced from their current basin of attraction into another one. The red, magenta and green bars represent the performances of the best (S), (C) and (CR) robots, respectively. Each bar represents the percentage of trials in which the robots were able to reach the appropriate destination despite the displacement. The "instant" condition refers to a situation in which immediately after the displacement the robots were allowed to move normally. The "1 second" and "3 seconds" conditions refer to a situation in which the robots were unable to move for 1 and 3 seconds, respectively. The "no-displacement" condition refers to a normal situation in which the robots were not subjected to displacements. During each trial the robot was displaced when it reached an imaginary line located 40cm before the first junction, 40cm after the first junction, in the middle of the second corridor or 40cm before the second junction. Data were collected from 2400 trials.

from the end of the first corridor on and used this information to move toward the appropriate destination and to preserve the relevant information (i.e. to maintain a specific relative position and orientation with respect to the environment). (C) and (CR) robots also exploited the possibility of integrating sensory-motor information over time to regulate their motor behaviour in the very first portion of the first corridor so as to reduce the variability with which they reached the green stimuli. Moreover, (C) and (CR) robots also exploited the possibility of integrating sensory-motor information over time to partially filter out the effect of noise affecting their sensors and motors and to better separate the trajectories produced while they navigated toward different target destinations.

The destination reached by (S) and (C) robots was determined by the position and the orientation assumed by the robots from the end of the first corridor on and was not affected by the type of stimuli experienced previously. Indeed, when these robots were displaced from their current position inside a certain basin of attraction into a position and an orientation located in a different basin of attraction, they navigated towards the target destination corresponding to the second basin of attraction in 98% of cases. The destination reached by (CR) robots, instead, also depended on the state of the internal neurons that encoded information about previously experienced sensory states. Indeed, displaced (CR) robots navigated toward the destination corresponding to the basin of attraction in which they were located before the displacement in 60% of cases. They managed to compensate the effect of the displacement by re-entering the basin of attractions in which they had been previously located. They navigated toward wrong destinations only in the remaining 40% of cases (see Figure 4.14).

4.4 Limiting cognitive offloading does not promote but rather prevents the synthesis of effective solutions

Here we report a series of experiments carried out to verify whether the discovery of sub-optimal reactive strategies that rely on cognitive offloading prevents the discovery of better cognitive solutions. In other words, we verify the hypothesis that cognitive

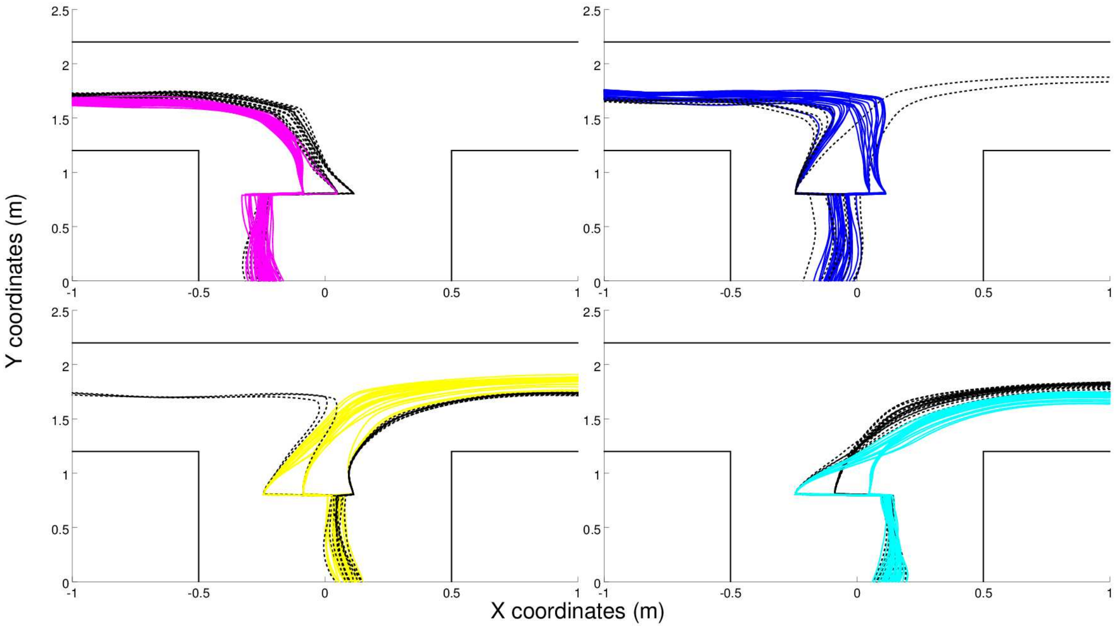


Figure 4.14: Trajectories displayed by the best (CR) robot at the first T-junction during a post-evaluation test in which the robots were displaced into one position and orientation located within one of the other three basins of attraction. The displacement was performed when the robot reached a distance of 40cm from the first junction. After the displacement the robot was allowed to move immediately (i.e. it was not blocked). Data were collected from over 200 trials, 50 for each combination of green stimuli. The colors indicate the target destinations (magenta: left-bottom, blue: left-top, cyan: right-bottom, and yellow: right-top). The black trajectories indicate the trials in which the robot reached a wrong destination, i.e. was unable to re-enter the correct basin of attraction after the displacement.

offloading constitutes a sort of shortcut that enables evolving individuals to improve their adaptive ability up to a certain level through the utilisation of solutions that are parsimonious from the perspective of the control system of the robot but that prevent the discovery of more complex and effective strategies. To achieve this objective we analysed the solutions found by evolving robots in situations in which the possibility of relying on cognitive offloading was reduced or prevented. One way to prevent the possibility of relying on cognitive offloading in the case of our experiments was to drastically reduce the width of the corridors. As we have seen, in fact, robots offload information concerning the type of green stimuli they have experienced by assuming different relative positions/orientations inside corridors. The utilisation of narrow corridors severely restricts the possibility of carrying out this type of offloading.

Analysis of the results obtained in a series of control experiments in which the width of the corridors was set to 29cm only indicates that, as expected, the use of highly constrained environmental conditions prevents the exploitation of cognitive offloading, i.e. the evolution of solutions analogous to that described in the previous section (results not shown). However, analysis of the performance of the robots evolved in this condition indicates that the elimination of solutions based on cognitive offloading does not lead to effective solutions. Indeed, it causes a drastic reduction of the robots' performance with respect to the normal condition (Figure 4.15). This implies that the elimination of solutions relying on cognitive offloading does not facilitate the evolution of alternative cognitive solutions. In other words, the lack of evolution of effective cognitive solutions cannot be explained simply by the availability of cognitive offloading "shortcuts".

Indeed, in most of the replications the best (Sn), (Cn) and (CRn) evolved robots were able to navigate correctly to only one of the four destinations and consequently succeeded in about one fourth of the trials. The evolved robots managed to navigate to one of the four correct destinations in most of the trials. Consequently, the impact of the errors occurred when the robots remained stuck, navigated erroneously back to-

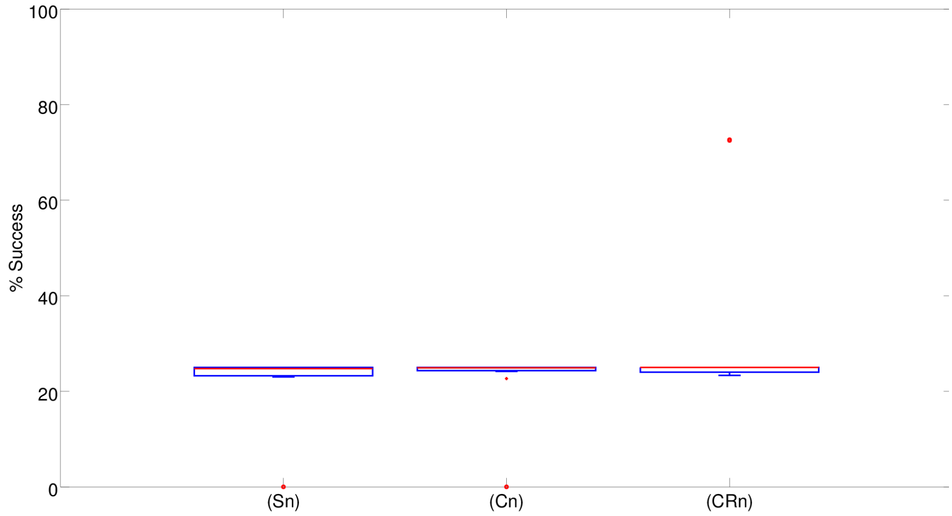


Figure 4.15: Performances obtained in the control experiment with narrow corridors. The boxplots show the percentage of successful trials carried out by the best 10 robots evolved in the (S), (C), and (CR) experimental conditions in 10 corresponding replications of the experiment. Boxes represent the inter-quartile range of the data. 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.

ward the central corridor or crashed into obstacles was marginal. We used (Sn), (Cn) and (CRn) to indicate the robots evolved in the narrow corridor condition. In only two replications, the best (CRn) robots managed to navigate correctly toward three out of the four destinations by succeeding in about three fourths of the trials (Figure 4.15). The performance of these two robots was similar to that achieved by the best (S) robot (Fisher exact test, $p=0.747$ and $p=0.797$), which could only rely on reactive strategies (Figure 4.7). However, it was significantly worse than that of the better (C) and (CR) robots, (Fisher exact test, $p<0.001$). Overall, the performance of the robots that evolved in these control experiments (Figure 4.15) was significantly worse than the performance obtained in the standard experiments (Figure 4.7) (Mann-Whitney U, $p<0.001$ for (S) and (C), and $p=0.01$ for (CR)).

Another mechanism that can be used to discourage evolving robots from relying on cognitive offloading is to randomly vary the position and orientation of the robots while they travel in the maze. To investigate the effect of this type of perturbation we carried out a series of control experiments in a standard maze in which every 50ms the position and the orientation of the robot were perturbed with a 15% probability. The perturbations were created by displacing the robot to the left or the right of a distance d and by varying the orientation of the robot by an angle a, where d and a were selected randomly with a uniform distribution within the range [3, 9]cm and [-15, 15]°, respectively. This type of perturbation drastically reduces the utility (usefulness) of offloading information in the relative positions and orientation of the robot in the environment. Once again, the hypothesis under verification is whether the introduction of a constraint that discourages the development of cognitive offloading solutions will favour the development of alternative, and possibly better, solutions. We used (Sp), (Cp) and (CRp) to indicate the robots evolved in the standard maze subjected to position and orientation perturbations during evolution.

The fact that this form of perturbation drastically reduces the usefulness of cognitive offloading is demonstrated by the fact that the performance of (Sp) robots, which can only rely on reactive strategies, drops to very low levels (Figure 4.16). The introduction of perturbations, however, causes a significant drop in performance with respect to the experiments without perturbations also in the case of (Cp) and (CRp) robots (Mann-Whitney U, $p<0.001$) (Figure 4.16). Notice that Figure 4.16 displays the results of a post-evaluation test in which the robots are evaluated in the absence of position and orientation perturbations.

Also in this case, therefore, the introduction of constraints that discourage the development of reactive solutions relying on cognitive offloading does not promote the evolution of effective cognitive solutions but rather prevents the possibility of synthesising good solutions of any kind.

As we will show in the next section, this negative result is not due to the impossibility of generating solutions that are able to compensate the effect of position and orientation perturbations. Indeed, evolving robots can solve the navigation problem in a close to optimal manner and can neutralise to a good extent the effect of position and orienta-

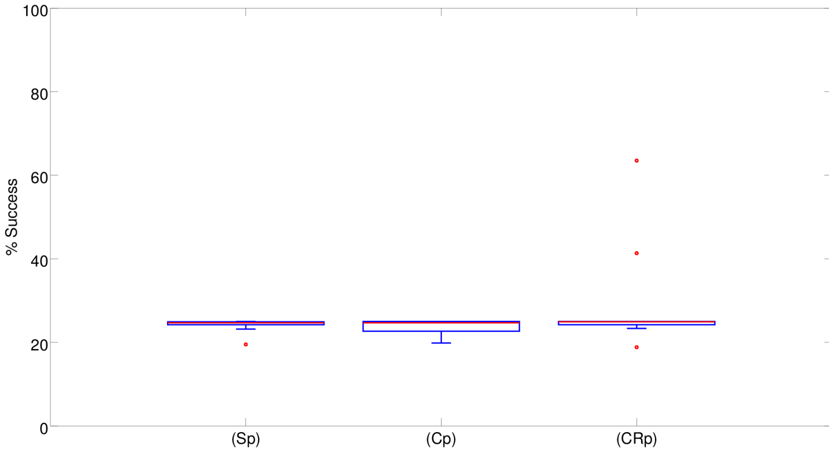


Figure 4.16: Performance obtained by robots that were subjected to position and orientation perturbations in all trials during 4000 generations. Data were obtained by post-evaluating the robots in a normal condition in which they were not subjected to perturbations. The boxplots show the percentage of successful trials carried out by the best 10 robots evolved in the (Sp), (Cp), and (CRp) experimental conditions in 10 corresponding replications of the experiment. Boxes represent the inter-quartile range of the data. 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.

tion perturbations, providing that the constraints which discourage the development of cognitive offloading strategies are not too strong.

4.4.1 The acquisition of reactive strategies promotes the evolution of cognitive capabilities

As demonstrated in the previous section, preventing or severely limiting the possibility of developing strategies based on cognitive offloading prevents the development of effective solutions. In this section, we demonstrate how the acquisition of reactive strategies promotes the evolution of cognitive solutions that enable the robots to accomplish their task also in conditions that cannot be mastered appropriately only by using cognitive offloading strategies.

To verify this hypothesis we carried out a new series of experiments in which we weakened the constraints that discourage the utilisation of cognitive offloading and increase the demand for cognitive solutions. This was carried out by subjecting the

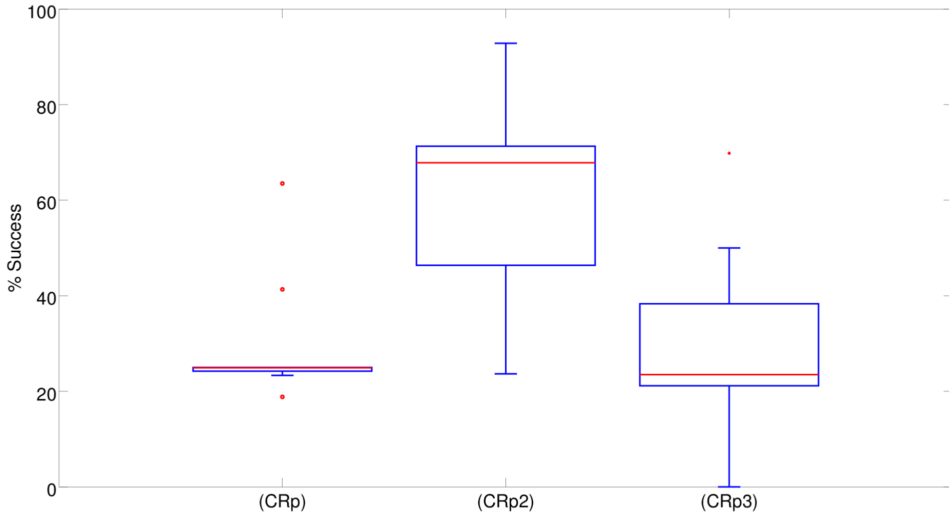


Figure 4.17: Performances obtained in experiments in which the robots were subjected to position and orientation perturbations in all cases (CRp), during half of the trials (CRp2) and during the second evolutionary phase (CRp3). Data obtained by post-evaluating the robots without position and orientation perturbations. The boxplots show the percentage of successful trials carried out by the best 10 robots evolved in each of the experimental conditions in 10 corresponding replications of the experiment. Boxes represent the inter-quartile range of the data. 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.

robots to position and orientation perturbations in only half of the trials (Figure 4.17, CRp2 condition) or by subjecting the robots to perturbations only during the second phase of the evolutionary process, i.e. from generation 2001 to 4000 (Figure 4.17, CRp3 condition). Notice that the Figure 4.17 displays the results of a post-evaluation test in which the evolved robots were not exposed to position and orientation perturbations.

The performances of the robots that were subjected to perturbations in only half of the trials are significantly better than the performances of the robots that experienced perturbations in all trials (Figure 4.17, CRp2 and CRp, Mann-Whitney U, p=0.004). The performances of the robots that were subjected to perturbations during the second phase of the evolutionary process, instead, did not differ significantly from those of the robots that were subjected to perturbations during all generations (Figure 4.17, CRp3 and CRp, Mann-Whitney U, p=0.36).

The fact that the robots evolved in the CRp2 condition relied on effective cognitive mechanisms can be demonstrated by post-evaluating the best robot in a control condition in which it was systematically displaced from its current position and orientation into another position and orientation located in a different basin of attraction through the same procedure described in Section 3. As can be seen, unlike the best CR robot, the best CRp2 robot is able to recover from the displacements in most cases, also in the condition in which it is blocked for three seconds after being displaced (Figure 4.18). The robot compensates for the effect of displacement by re-entering into the basin of attraction in which it was located before the displacement. This is carried out by storing in its internal states information encoding the type of basin of attraction in which it is located and by using this information to re-enter into the previous basin of attraction, as shown in Figure 4.14. The best CR and CRp2 robots make use of the internal states to neutralise the effects of this type of displacement (see Figures 4.13 and 4.18). However, the best CRp2 displays much better ability. This is not surprising, because CRp2 robots were subjected to position and orientation perturbations during evolution. In normal conditions, i.e. without displacements or perturbations, the performances of the best CR and CRp2 controllers do not differ statistically (Fisher Exact Test, p=0.308).

The fact that the development of cognitive offloading strategies supports the development of effective cognitive abilities, such as those displayed by the best CRp2 robot, is also demonstrated by the analysis of the course of the evolutionary process in the case of the best replication of the experiment. In fact, the post-evaluation of the best robots in every 50 generations indicates that the development of the ability to master the trials not affected by displacements precedes the development of the ability to also master the trials subjected to displacement (Figure 4.19). The performances in the two conditions differed significantly from generation 1500 to 1800 with the exception of generation 1700 (Fisher Exact Test, $p<0.05$). We focused on this 500 generations

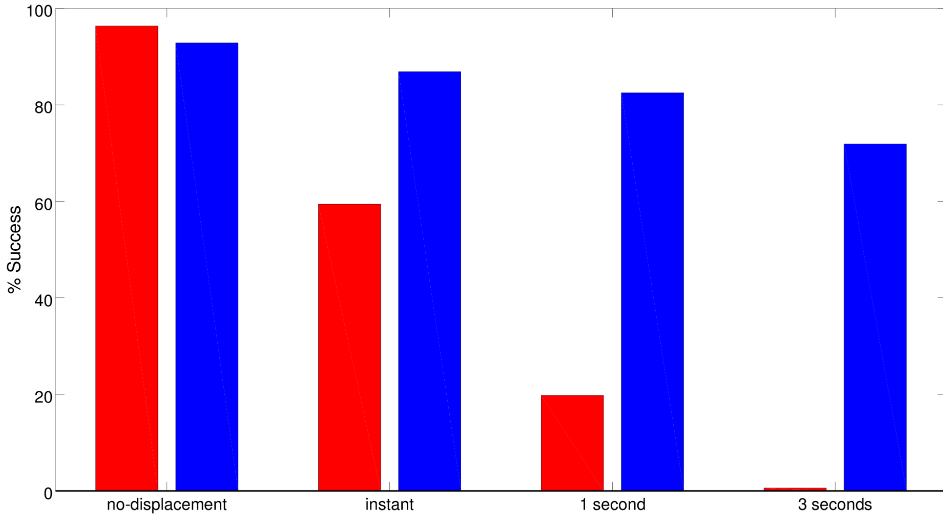


Figure 4.18: Performance displayed by the best (CR) and (CRp2) robots during a post-evaluation test in which the robots were systematically displaced into a position and orientation located in a different basin of attraction (red and blue bars, respectively). Each bar represents the percentage of trials in which the robots were able to reach the appropriate destination despite the displacement. The “instant” condition refers to a situation in which the robots were allowed to move normally immediately after the displacement. The “1 second” and “3 seconds” conditions refer to situations in which the robots were blocked for 1 and 3 seconds, respectively, after the displacement. The “no-displacement” condition refers to a normal situation in which the robots were not subjected to displacements. The robots were displaced once during each trial. The displacement occurred when they reached an imaginary line located 40cm before the first junction, or 40cm after the first junction, or in the middle of the horizontal corridor, or 40cm before the second junction. Data collected in 2400 trials.

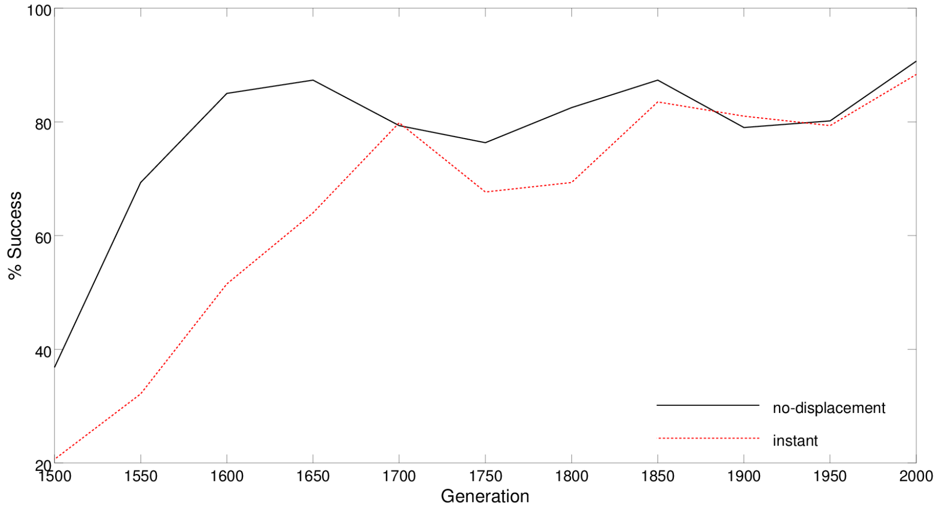


Figure 4.19: Performance displayed by the best CRp2 during the course of the evolutionary process in a no-displaced and displaced condition (black and red lines, respectively). In the latter condition the robot was allowed to move immediately after the displacement. Data were collected every 50 generations from generation 1500 to 2000 and averaged over 600 trials in the case of the no-displacement condition and 2400 trials in the case of the displaced condition. The displacement occurs when the robot reaches an imaginary line located 40cm before the first junction, or 40cm after the first junction, or in the middle of the horizontal corridor, or 40cm before the second junction.

period because in the case of the best replication this is the phase in which most progress occurs. Thus, the development of a strategy that operates primarily on the basis of cognitive offloading and that enables the robots to handle the navigation task in the normal condition but not in the condition with displacements supports the development of a hybrid strategy that relies also on internal states and that enables the robots to also master the trials affected by displacements.

Incidentally, the internal neurons of the (CR) and (CRp2) robots are somewhat similar to the hippocampal cells of rodents located in maze environments that encode information about the location of the animal in the maze and about the trajectory the animal is performing and will perform to reach the target destination [136, 187, 191]. However, a detailed analysis of the relation with these neurophysiological findings is outside the scope of this study.

At this point we will try to explain why evolving robots: (i) always develop cognitive offloading strategies, (ii) are unable to develop effective strategies relying exclusively on internal states even in conditions in which the possibility of using cognitive offloading is completely ruled out, (iii) are able to extract internal states encoding target destinations that are maintained over time and are used to compensate the effects of position and orientation perturbations.

We believe that to explain these results we have to appreciate the full complexity of the task. In particular, we must consider that the problem involves the ability to both travel in the maze environment by avoiding collisions and inversions in the direction of motion and the ability to turn in the appropriate direction at junction areas. The most straightforward way of navigating in the maze and avoiding collisions is to use a reactive strategy that regulates the direction of the robot on the basis of the current state of the infrared sensors. In other words, the ability to navigate in the maze environment necessarily requires use of a reactive strategy. The capacity to turn in the appropriate direction at junction areas, instead, can be obtained by either exploiting primarily external cues generated through cognitive offloading or exploiting primarily internal cues generated by integrating sensory-motor information over time. Using a mixed strategy that operates on the basis of both reactive rules for the purpose of navigation and more complex rules for the purpose of decision-making at junctions necessarily requires the incorporation of mechanisms that sort out the conflicts arising between reactive and non-reactive control rules.

As an example of conflict, let us consider the case of a robot that decides which direction to take at a junctions on the basis of its internal state. Moreover, let us imagine that the robot is located at the beginning of the first junction and that it is traveling toward one of the two left destinations. During the first and the third part of the junction negotiation the robot should not turn too sharply left to avoid colliding with the left walls. During the central part of the junction negotiation, instead, the robot should definitely turn left following the indication coming from its internal state. The conflict arising between obstacle avoidance and the decision-making rules that control the direction of turns at junctions could be solved by regulating the relative importance of the two rules on the basis of the relative position of the robot within the junction. However, the robot does not have access to position information. It only perceives the proximity of nearby obstacles. This implies that the robot would be forced to determine the relative importance of the two alternative control rules on the basis of the indirect, noisy, and incomplete information provided by its sensors. Consequently, this implies that the probability that the robot will fail as a result of ineffective regulation is not negligible.

Instead, the behavioural attractor strategies displayed by evolved robots do not require differentiating the way in which the robots react to sensory states experienced in different portions of the environment. Indeed, once the robots enter into the right basin of attraction, they just need to keep moving on the basis of simple reactive rules in both corridors and junctions. One reason why evolving robots always select cognitive offloading solutions is that these strategies are more robust and less prone to errors with respect to strategies in which turning decisions are made on the basis of internal states.

The second reason why evolution always converges toward cognitive offloading strategies is that preparatory actions that anticipate in part the execution of the required behaviour are adaptive and therefore tend to be selected. This implies that the individuals that anticipate the movement toward the left or the right side of the corridor during the trials in which they should turn toward that side at the first and/or at the second junctions tend to be selected. This produces a progressive anticipation of the time when the turning actions are initiated that ultimately leads to a situation in which they are initiated up to the point when the robot perceives the green stimuli. This in turn eliminates the need to extract and use internal states that encode information about previously experienced stimuli. Indeed, we might say that in the evolved robots the left or right turning behaviours produced at the first and second junctions are initiated already during the first half of the first corridor, when the robot perceives the position of the green stimuli. In other words, by anticipating action execution through preparatory actions, the robots manage to transform a time delay task into a simpler problem that does not include any temporal offset between the perception of stimuli and the initiation of the action afforded by the stimuli. Overall, this implies that the selection of cognitive offloading strategies of this type is inevitable, at least in the case of the double T-Maze experimental setting.

The tendency to anticipate behaviours through the execution of preparatory actions provides two advantages: (i) it enables the execution of smoother transitions between behaviours (i.e. between the navigation behaviour performed within the corridors and the turning behaviour performed within the junctions), and (ii) it enables reducing and/or eliminating the time delay between the moment when the stimuli affording a given behaviour are experienced and the moment when the behaviour afforded by the stimuli is executed.

As we have showed above, however, the ability to solve the time-delay problem through preparatory actions does not necessarily prevent development of the ability to extract information encoding the type of basin of attraction in which the robot is currently located or the green stimuli that the robot experienced and the development of an ability to use these internal states to navigate to the appropriate destination. Indeed, as we have seen, (CR) robots display the ability to re-enter into the basin of attraction in which they were previously located after being displaced into another wrong basin of attraction. In the case of robots subjected to a moderate level of position and orientation perturbation (CRp2), this cognitive capability is so effective that it enables the robots to compensate for the effect of the displacement in most cases, even when the robots are blocked for three seconds after the displacement. This type of redundant solution enables these robots to exploit the advantage of preparatory actions and avoid the problems caused by noise and by position perturbations.

Anticipation is a widespread phenomenon in sequential motor control. The preparatory actions that support the realisation of effective grasping behaviours are an example of this. These preparatory actions involve appropriate modification of the posture of the hand performed during execution of the reaching behaviour that precedes the grasping action [181]. A second example is constituted by the co-articulatory movements produced by sign language interpreters engaged in fingerspelling. Indeed, the posture of the hand used to indicate a letter is influenced by the posture that the hand should assume later to indicate the following letters [81].

4.5 Discussion

The present chapter presented the results and analysis that clarify whether the development of reactive solutions promotes or prevents the evolution of cognitive robots in problems in which reactive control policies enable the achievement of sub-optimal performance only. For this purpose a series of experiments were carried out in which evolving robots had to solve a time-delay task in a double T-maze environment in which the destination to be reached depended on the stimuli perceived by the robot during the initial phase of the navigation. The problem chosen is qualitatively similar but more complex than the tasks used in previous studies that investigated the evolution of cognitive capabilities [4, 9, 29, 41, 99, 147]. The additional complexity is that the task used here involves a greater number of different destinations, requires making two subsequent stimuli-dependent decisions, involves a longer time delay between the moment in which the robot experiences the stimuli and the moment in which it should make the corresponding turning decisions, and involves variations affecting the initial position and orientation of the robot, the position of the stimuli and the size of the environment.

Analysis of the experiment in which the robots were provided with reactive controllers confirmed that the problem does not allow for optimal, or close-to-optimal, reactive solutions (Figure 4.6). Surprisingly, however, reactive robots managed to solve the task to a large extent. Analysis of the solutions discovered by the evolving robots indicate that this is achieved by exploiting cognitive offloading. Indeed, the evolved robots display an ability to extract critical states, store these states in the robot/environmental relations and regulate their behaviour on the basis of their relative position in the environment (Figure 4.7).

Analysis of the robots provided with richer neural controllers indicated that the possibility of storing internal states enables the evolving robots to achieve close-to-optimal performance, i.e. to achieve better performance with respect to reactive robots (Figure 4.6). Analysis of these experiments indicates that cognitive offloading also plays a key role in these robots (Figure 4.11-4.13). The achievement of better results is due to the development of additional cognitive capabilities that overcome the limitations displayed by robots that operate on the basis of reactive controllers only. The results obtained thus indicate that the development of the reactive strategies based on cognitive offloading does not prevent the development of solutions that rely on cognitive capabilities.

This conclusion is further supported by results obtained in other experiments in which the usefulness of cognitive offloading was reduced or eliminated by using an environment formed by narrow corridors or by subjecting evolving robots to position and orientation perturbations. As expected, the robots evolved in these conditions relied less or not at all on cognitive offloading. This, however, did not enable the robots to discover alternative strategies for solving the task. This simply led to the evolution of robots displaying rather ineffective solutions (Figures 4.15 and 4.16). Overall, this indicates that the elimination of cognitive offloading does not promote but rather prevents the synthesis of effective solutions.

Finally, the results obtained demonstrated how the acquisition of reactive strategies promotes the evolution of cognitive strategies, or better of hybrid strategies including both cognitive offloading and cognitive mechanisms. These hybrid strategies enable the robots to navigate toward the appropriate destination also after being displaced into another basin of attraction and also after being blocked there for three seconds (Figures 4.17 and 4.18). This type of solution was obtained by weakening the constraints that reduce the usefulness of cognitive offloading (i.e. by perturbing the position and orientation of the robot in only half of the trials). This, indeed, creates the appropriate demand for the development of cognitive abilities without preventing the development of cognitive offloading strategies.

Overall, these results indicate that reactive strategies relying on cognitive offloading do not necessarily constitute a dead end that might retard or prevent the evolution of better cognitive strategies. On the contrary, they constitute an important component of effective solutions and can co-exist and support the development of complementary cognitive capabilities. For results collected in human subjects that indicate how cognitive offloading can favour the acquisition of abstract concepts see [174].

The importance of the incremental nature of the evolutionary process and of the acquisition of reactive strategies in the synthesis of better cognitive strategies is further demonstrated by the analysis of the course of the evolutionary process. Indeed, the evolution of the cognitive offloading ability precedes the evolution of the cognitive capabilities that use internal states to determine the travel destination (Figure 4.19).

One question that remains open is whether solutions that are purely cognitive, i.e. that do not rely also on cognitive offloading, exist and can be discovered. It is worth noting that the existence of this type of solution cannot be taken for granted, at least in the context of the domain considered in this study.

In general, analysis of the characteristics of evolved strategies suggests that reactive and cognitive components of control policies should not be considered as neatly separable entities performing well-differentiated functions. Regulation of the robot's actions performed on the basis of internal states should always be integrated with regulation of the robot's actions, which is carried out on the basis of currently perceived states. In the context of the presented experiments this implies that the efficacy of a cognitive component of the robot policy that determines the turning direction at T-junctions on the basis of internal states strongly depends on the way the robot reacts to perceived stimuli independently of the value of its internal states, and vice versa. Moreover, the use of a cognitive offloading strategy for the purpose of storing information about previously experienced stimuli does not necessarily conflict with use of internal states that have the same function. On the contrary, the combined use of two alternative mechanisms with different characteristics for achieving the same function might provide advantages. If we embrace a less simplified view of the relation between reactive and cognitive components we see fewer reasons to expect interferences and more reasons to expect synergies, like those we found in the presented experiments.

Robotics is a field that has already changed our world substantially with industrial robotics, and that still has the potential to change it even more with service and mobile robots. However, operating on everyday life environments poses many challenges due to the fact that they are highly dynamic and uncertain environments. So, dealing with these challenges requires a research effort for developing and understanding the innovative methods that could embed artificial embodied agents, aka robots, with some sort of intelligence for being adaptive.

A key aspect for mobile robots dealing with complex environments is the capacity of displaying adaptive behaviours. It can enable robots to achieve the required function(s) and display a behavioural organization that facilitates the possibility to adapt to varying environmental conditions. Particularly some interesting types of adaptive behaviour that seem to bring these benefits are the capability to rely on multiple behavioural and cognitive capabilities in an integrated manner. As presented in the first chapter, adaptive behaviours can originate from the exploitation of environmental complexity, behaviours composed by interconnected and interdependent subunits, aka articulated behaviours, and also from the capacity of using internal information integrated over time in combination with external stimuli, aka cognitive behaviours.

Among the different approaches that could contribute to the development of adaptive robots are those relying on artificial evolution, which can synthesise not only robot controllers but also body-plans. These methods allow the so called design by emergence [135], the possibility to exploit behavioural properties that originate from the interaction among environment, body and mind. However, the outcome of the evolutionary process depends on the characteristics of the agents, and the environment in which the agents evolve, and these are determined by the experimenter. Understanding how these characteristics should be set so to maximise the results of evolution is far from trivial.

In this thesis I analyzed some of the aspects that can promote the evolution of adaptive behaviour, i.e. behaviour that permit to achieve certain functions but that also present an organization that facilitates the adaptation to varying environmental conditions, and the exploitation of different capabilities in an integrated manner. The results presented showed how behavioural plasticity and the development of integrated behavioural and cognitive capabilities can enable evolving robots to achieve better solutions. Moreover I analyzed the mechanisms that can promote the development of plastic behaviours and the development of robots that operate on the basis of integrated behavioural and cognitive capabilities. For example I clarified the importance of affordance generation and the sensory-motor and neural characteristics that can promote affordance generation.

The kind of behavioural plasticity evolved and analysed in the chapter 2 differs from neural plasticity in which there is synaptic weights adjustment for modifying the overall behaviour. It differs also from the approaches that structurally divide the robot controllers in order to explicitly produce articulated behaviours. As confirmed by the results presented, the benefits of this functional type of modularity is that the connectedness among behaviours produces smooth and effective behavioural integration. This overall produces more effective behaviours compared to non-plastic behaviours, or structurally modular behaviours in which there is a sharp switch. We saw that the main mechanism behind the evolution and production of the behavioural plasticity was the affordance generation. This mechanism allowed the combination of different type of stimuli, internal and external, to arbitrate and regulate the execution of the different behaviours, and to precisely define the switch point among them.

Regarding the evolution of cognitive agents, i.e. those capable of integrating information over time and using it for decision making, we saw that the use of reactive strategies, more specifically those that offload part of the cognitive process to the environment, can promote cognitive development rather than prevent it. Furthermore, we saw that this kind of strategy can be combined with cognitive solutions, those relying on internal states, in order to compose an effective behaviour. According to the results presented in chapter 3, encoding part of the information and offloading part of the cognitive work into the external environment can facilitate the development of cognitive agents. This is confirmed by the fact that the evolution of cognitive offloading precedes the development of strategies that actively use internal information for regulating the behaviour. And this happens even in the experiments in which there is a stronger pressure for selecting individuals that rely more on the use of internal information, as in the experiment with a moderate level of position noise. Furthermore, when the possibility to evolve cognitive offloading was removed, narrow corridors or extreme position noise, no effective solution was found by the evolutionary process.

While in the literature it is common to find studies investigating how to develop effective solutions using artificial evolution through changes in the fitness function [71, 95, 151], or in the structure of the genetic algorithm [21, 138, 141], enlightening specific aspects of the artificial evolution that can provide an evolvability increase regarding the development of adaptive and effective behaviours is a less explored direction, and the one chosen in this thesis. Building knowledge about these aspects can represent a promising step for taking Evolutionary Robotics approaches to the next level, i.e. effectively creating robots capable of operating and providing services in the real world. Additionaly, the inspiration on evolutionary biology and ethology may not only provide insights and help the guidance on which research directions to follow, but also in the other way around, provide a small contribution also to those fields, allowing to check if artificial agents are capable of presenting characteristics observed in biological ones, and also to compare and study the resulting mechanisms.

4.6 Contribution to knowledge

Summarising the contributions to knowledge provided by this thesis, we can say that:

- the studies presented in this thesis show that there is an alternative and quite unexplored way, other than focusing on algorithms, fitness functions or particular types of robot controller, for generating adaptive robotic behaviour. The studies presented here shed some light in important aspects of evolutionary processes that can lead to the evolution of adaptive behaviours, more specifically those regarding articulated and cognitive ones, without proposing any modification in the evolutionary process that could create pressure or bias the development of these types of behaviours.

- the results and analysis presented on chapter 2 show how behavioural plastic robots can evolve and effectively perform even in a single goal task;
- behavioural plasticity can be evolved even without the use of synaptic plasticity or structural division of the robot controllers;
- the lack of a sharp division among modules, allows a high level of interconnectedness leading to smooth and effective transitions among behaviours;
- in a behavioural plastic strategy, the possibility of arbitrating the sub-behaviours based on multiple stimuli, internal and external, allow smoother and more adaptive switches, i.e. the transitions happen in the precise time leading to the most effective task performance;
- the possibility to generate affordances for maintaining or switching sub-behaviours, through the robot's action, represents an important mechanism that enables the evolution of behavioural plasticity in evolving robots;
- the results and analysis presented on chapter 3 clarified the important role that reactive strategies, specifically cognitive offloading, have in the evolution and operation of minimally cognitive robots;
- contrary to what some authors claim, that cognitive offloading could prevent the development of cognitive capabilities, these strategies may in fact precede the development of strategies relying on internal information, and the possibility of using cognitive offloading may promote the evolution of cognitive capabilities.
- the best cognitive solutions found are actually a combination of cognitive offloading and the use of internal information integrated over time. These strategies are complementary, i.e., internal states are reinforced by the stimuli differentiation

created by the cognitive offloading strategy, conversely, the internal information is used to correct disturbances and maintain the robot in the proper attractors created by the reactive strategy. This overall suggests that reactive and cognitive strategies can be less separable than one could initially think, instead, they may work together in the composition of a hybrid superior solution as was the case of the results presented in this work.

- it is possible to bias the evolution to create solutions that rely more on internal information than on external stimuli by reducing the reliability of external information, adding position noise for instance, but these solutions may not bring any performance advantage.

References

[1] Evan Ackerman. Long exposure pictures of robots cleaning. `http://spectrum.ieee.org/automaton/robotics/home-robots/long-exposure-pictures-of-robots-cleaning`, oct 2010. Acessed: 2015-04-23.

[2] Michael L. Anderson. Embodied Cognition: A field guide. *Artificial Intelligence*, 149(1):91–130, September 2003.

[3] R. J. Anderson. Building a modular robot control system using passivity and scattering theory. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 1, pages 698–705 vol.1, April 1996.

[4] E. Antonelo, B. Schrauwen, and D. Stroobandt. Mobile robot control in the road sign problem using Reservoir Computing networks. In *IEEE International Conference on Robotics and Automation, 2008. ICRA 2008*, pages 911–916, May 2008.

[5] Brenna D. Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, May 2009.

[6] R. Arkin. Motor schema based navigation for a mobile robot: An approach to programming by behavior. In *1987 IEEE International Conference on Robotics and Automation Proceedings*, volume 4, pages 264–271, March 1987.

[7] Ronald C. Arkin. Motor Schema — Based Mobile Robot Navigation. *The International Journal of Robotics Research*, 8(4):92–112, August 1989.

[8] Andrea Avella and Dinesh K Pai. Modularity for sensorimotor control: evidence and a new prediction. *Journal of motor behavior*, 42(6):361–369, 2010.

[9] B. Bakker, V. Zhumatiy, G. Gruener, and J. Schmidhuber. A robot that reinforcement-learns to identify and memorize important previous observations. In *2003 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings*, volume 1, pages 430–435 vol.1, October 2003.

[10] G. J. Barlow, C. K. Oh, and E. Grant. Incremental evolution of autonomous controllers for unmanned aerial vehicles using multi-objective genetic programming. In *IEEE Conference on Cybernetics and Intelligent Systems, 2004.*, volume 2, pages 689–694, December 2004.

[11] G. W. Barlow. Modal Action Patterns. In *In T. A. Sebeok (ed.), How Animals Communicate.* Indiana University Press, Bloomigton:, 1977.

[12] C. J. Barnard. *Animal Behaviour: Ecology and Evolution*. Springer Science & Business Media, December 1983.

[13] Randall D. Beer. Toward the Evolution of Dynamical Neural Networks for Minimally Cognitive Behavior. In *From Animals to Animats 4: Proc. 4th Int. Conf. on Simulation of Adaptive Behavior*, pages pp. 421–429. MIT Press, 1996.

[14] Randall D. Beer. The Dynamics of Active Categorical Perception in an Evolved Model Agent. *Adaptive Behavior*, 11(4):209–243, December 2003.

[15] Randall D. Beer and John C. Gallagher. Evolving Dynamical Neural Networks for Adaptive Behavior. *Adaptive Behavior*, 1(1):91–122, June 1992.

[16] Randall D. Beer and Paul L. Williams. Information Processing and Dynamics in Minimally Cognitive Agents. *Cognitive Science*, 39(1):1–38, January 2015.

[17] Raffaele Bianco and Stefano Nolfi. Toward open-ended evolutionary robotics: evolving elementary robotic units able to self-assemble and self-reproduce. *Connection Science*, 16(4):227–248, December 2004.

[18] E. Bizzi, S. F. Giszter, E. Loeb, F. A. Mussa-Ivaldi, and P. Saltiel. Modular or-

ganization of motor behavior in the frog's spinal cord. *Trends in Neurosciences*, 18(10):442–446, October 1995.

[19] Jesper Blynel and Dario Floreano. Exploring the T-Maze: Evolving Learning-Like Robot Behaviors Using CTRNNs. In Stefano Cagnoni, Colin G. Johnson, Juan J. Romero Cardalda, Elena Marchiori, David W. Corne, Jean-Arcady Meyer, Jens Gottlieb, Martin Middendorf, Agnès Guillot, Günther R. Raidl, and Emma Hart, editors, *Applications of Evolutionary Computing*, number 2611 in Lecture Notes in Computer Science, pages 593–604. Springer Berlin Heidelberg, April 2003.

[20] M. Bonani, V. Longchamp, S. Magnenat, P. Retornaz, D. Burnier, G. Roulet, F. Vaussard, H. Bleuler, and F. Mondada. 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 (IROS)*, pages 4187–4193, October 2010.

[21] Josh Bongard, Paul Beliveau, and Gregory Hornby. Avoiding local optima with interactive evolutionary robotics. In *Proceedings of the 14th annual conference companion on Genetic and evolutionary computation*, pages 1405–1406. ACM, 2012.

[22] Josh C. Bongard. Spontaneous Evolution of Structural Modularity in Robot Neural Network Controllers. In *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, GECCO '11, pages 251–258, New York, NY, USA, 2011. ACM.

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

[24] B Browning, J Bruce, M Bowling, and M Veloso. STP: Skills, tactics, and plays for multi-robot control in adversarial environments. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 219(1):33–52, February 2005.

[25] Raffaele Calabretta, Stefano Nolfi, Domenico Parisi, and Günter P. Wagner. Duplication of Modules Facilitates the Evolution of Functional Specialization. *Artificial Life*, 6(1):69–84, January 2000.

[26] Angelo Cangelosi, Josh Bongard, Martin H. Fischer, and Stefano Nolfi. Embodied Intelligence. In *Springer Handbook of Computational Intelligence*, pages 697–714. Springer, Berlin, Heidelberg, 2015. DOI: 10.1007/978-3-662-43505-2_37.

[27] Anthony Chemero. *Radical embodied cognitive science*. A Bradford Book. MIT Press, Cambridge, Mass., 2011. OCLC: 698839160.

[28] Hillel J. Chiel and Randall D. Beer. The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, 20(12):553–557, December 1997.

[29] Ji Ryang Chung and Yoonsuck Choe. Emergence of Memory in Reactive Agents Equipped With Environmental Markers. *IEEE Transactions on Autonomous Mental Development*, 3(3):257–271, September 2011.

[30] Andy Clark. *Being There: Putting Brain, Body, and World Together Again*. MIT Press, January 1997.

[31] Andy Clark. An embodied cognitive science? *Trends in Cognitive Sciences*, 3(9):345–351, September 1999.

[32] Andy Clark and David Chalmers. The extended mind. *Analysis*, 58(1):7–19, 1998.

[33] Steven H. Collins, Martijn Wisse, and Andy Ruina. A Three-Dimensional Passive-Dynamic Walking Robot with Two Legs and Knees. *The International Journal of Robotics Research*, 20(7):607–615, July 2001.

[34] M. Conrad. The geometry of evolution. *Bio Systems*, 24(1):61–81, 1990.

[35] H. Cruse, Ch. Bartling, M. Dreifert, J. Schmitz, D.E. Brunn, J. Dean, and T. Kindermann. Walking: A Complex Behavior Controlled by Simple Networks. *Adaptive Behavior*, 3(4):385–418, March 1995.

[36] A. Cully and J.-B. Mouret. Evolving a Behavioral Repertoire for a Walking Robot. *Evolutionary Computation*, 24(1):59–88, 2016.

[37] John Daintith and Elizabeth A. Martin. *A Dictionary of Science*. Oxford University Press, 2005.

[38] Ralf Der and Georg Martius. *The Playful Machine: Theoretical Foundation and Practical Realization of Self-Organizing Robots*. Springer Science & Business Media, January 2012.

[39] Thomas J. DeWitt and Samuel M. Scheiner. *Phenotypic Plasticity: Functional and Conceptual Approaches*. Oxford University Press, January 2004.

[40] Stephane Doncieux and Jean-Baptiste Mouret. Beyond black-box optimization: a review of selective pressures for evolutionary robotics. *Evolutionary Intelligence*, 7(2):71–93, July 2014.

[41] Miguel Duarte, Sancho Moura Oliveira, and Anders Lyhne Christensen. Evolution of Hybrid Robotic Controllers for Complex Tasks. *Journal of Intelligent & Robotic Systems*, 78(3-4):463–484, July 2014.

[42] S. Ducatez, J. N. Audet, and L. Lefebvre. Independent appearance of an innovative feeding behaviour in Antillean bullfinches. *Animal Cognition*, 16(3):525–529, May 2013.

[43] Lee Alan Dugatkin. *Principles of Animal Behavior*. W. W. Norton & Company, New York, 3 edition edition, March 2013.

[44] William G. Eberhard. Behavioral flexibility in orb web construction : Effects of supplies in different silk glands and spider size and weight. *Journal of Arachnology*, 16(3):295–302, 1988.

[45] A. Faíña, L. T. Jacobsen, and S. Risi. Automating the Incremental Evolution of Controllers for Physical Robots. *Artificial Life*, 23(2):142–168, May 2017.

[46] J.C. Fentress. A View of Ontogeny. *In: Esenberg J, Kleiman D (eds) Special publications American Society of Mammalogists*, 7:24–64, 1983.

[47] Paolo Fiorini and Erwin Prassler. Cleaning and Household Robots: A Technology Survey. *Autonomous Robots*, 9(3):227–235, December 2000.

[48] D. Floreano and F. Mondada. Evolution of Homing Navigation in a Real Mobile Robot. *IEEE Transactions on Systems, Man and Cybernetics Part B : Cybernetics*, 26(3):396–407, 1996.

[49] Dario Floreano, Peter Dürr, and Claudio Mattiussi. Neuroevolution: from architectures to learning. *Evolutionary Intelligence*, 1(1):47–62, March 2008.

[50] Dario Floreano, Phil Husbands, and Stefano Nolfi. Evolutionary Robotics. In Bruno Siciliano Prof and Oussama Khatib Prof, editors, *Springer Handbook of Robotics*, pages 1423–1451. Springer Berlin Heidelberg, 2008. DOI: 10.1007/978-3-540-30301-5_62.

[51] Dario Floreano and Stefano Nolfi. Adaptive behavior in competing co-evolving species. In *4th European Conference on Artificial Life*, pages 378–387, 1997.

[52] Dario Floreano and Joseba Urzelai. Neural Morphogenesis, Synaptic Plasticity, and Evolution. *Theory in Biosciences*, 120(3):225–240, January 2001.

[53] C. R. Gallistel. *The Organization of Action: A New Synthesis*. L. Erlbaum Associates, 1980.

[54] Francesca Garbarini and Mauro Adenzato. At the root of embodied cognition: Cognitive science meets neurophysiology. *Brain and Cognition*, 56(1):100–106, October 2004.

[55] James Jerome Gibson. *The ecological approach to visual perception*. Houghton Mifflin, 1979.

[56] Onofrio Gigliotta and Stefano Nolfi. On the Coupling Between Agent Internal and Agent/ Environmental Dynamics: Development of Spatial Representations in Evolving Autonomous Robots. *Adaptive Behavior - Animals, Animats, Software Agents, Robots, Adaptive Systems*, 16(2-3):148–165, 2008.

[57] Onofrio Gigliotta, Giovanni Pezzulo, and Stefano Nolfi. Evolution of a predictive internal model in an embodied and situated agent. *Theory in Biosciences = Theorie in Den Biowissenschaften*, 130(4):259–276, December 2011.

[58] Sam J. Gilbert. Strategic offloading of delayed intentions into the external environment. *Quarterly Journal of Experimental Psychology (2006)*, 68(5):971–992, 2015.

[59] Sam J. Gilbert. Strategic use of reminders: Influence of both domain-general and task-specific metacognitive confidence, independent of objective memory ability. *Consciousness and Cognition*, 33:245–260, May 2015.

[60] F. Montes Gonzalez, Tony J. Prescott, Kevin Gurney, Mark Humphries, and Peter Redgrave. An embodied model of action selection mechanisms in the vertebrate brain. *From animals to animats*, 6:157–166, 2000.

[61] Goren Gordon, Ehud Fonio, and Ehud Ahissar. Emergent exploration via novelty management. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 34(38):12646–12661, September 2014.

[62] Joachim de Greeff and Stefano Nolfi. Evolution of Implicit and Explicit Communication in Mobile Robots. In Stefano Nolfi and Marco Mirolli, editors, *Evolution of Communication and Language in Embodied Agents*, pages 179–214. Springer Berlin Heidelberg, 2010.

[63] Eduard Grinke, Christian Tetzlaff, Florentin Wörgötter, and Poramate Manoonpong. Synaptic plasticity in a recurrent neural network for versatile and adaptive behaviors of a walking robot. *Frontiers in Neurorobotics*, 9, 2015.

[64] Karin Gross, Gilberto Pasinelli, and Hansjoerg P. Kunc. Behavioral Plasticity Allows Short-Term Adjustment to a Novel Environment. *The American Naturalist*, 176(4):456–464, October 2010.

[65] C. Hartland, N. Bredeche, and M. Sebag. Memory-enhanced Evolutionary Robotics: The Echo State Network Approach. In *IEEE Congress on Evolutionary Computation, 2009. CEC '09*, pages 2788–2795, May 2009.

[66] Masahiko Haruno, Daniel M. Wolpert, and Mitsuo Kawato. Mosaic model for sensorimotor learning and control. *Neural computation*, 13(10):2201–2220, 2001.

[67] I. Harvey, P. Husbands, D. Cliff, A. Thompson, and N. Jakobi. Evolutionary robotics: the Sussex approach. *Robotics and Autonomous Systems*, 20(2):205–224, June 1997.

[68] Wade N. Hazel. The environmental and genetic control of seasonal polyphenism in larval color and its adaptive significance in a swallowtail butterfly. *Evolution; International Journal of Organic Evolution*, 56(2):342–348, February 2002.

[69] Robert A. Hinde. *Animal behavior: a synthesis of ethology and comparative psychology*. McGraw-Hill, New York, NY, US, 1966.

[70] John H. Holland. Genetic Algorithms and Adaptation. In *Adaptive Control of Ill-Defined Systems*, NATO Conference Series, pages 317–333. Springer, Boston, MA, 1984. DOI: 10.1007/978-1-4684-8941-5_21.

[71] N. van Hoorn, J. Togelius, and J. Schmidhuber. Hierarchical controller learning in a First-Person Shooter. In *2009 IEEE Symposium on Computational Intelligence and Games*, pages 294–301, September 2009.

[72] Steven Horst. The computational theory of mind. *Stanford Encyclopedia of Philosophy*. `https://stanford.library.sydney.edu.au/archives/win2011/entries/computational-mind/`, 2009. Online. Acessed: 2017-07.

[73] Joost Huizinga, Jeff Clune, and Jean-Baptiste Mouret. Evolving neural networks that are both modular and regular: HyperNEAT plus the connection cost technique. In *In: Proceedings of the genetic and evolutionary computation conference (GECCO)*, pages 697–704, New York, 2014. ACM Press.

[74] P. Husbands, I. Harvey, D. Cliff, and G. Miller. Artificial evolution: a new path for artificial intelligence? *Brain and Cognition*, 34(1), June 1997.

[75] iRobot. Our history. `http://www.irobot.com/about-irobot/company-information/history.aspx`, 2013. Accessed: 2015-04-29.

[76] I. Iturrate, L. Montesano, and J. Minguez. Robot reinforcement learning using EEG-based reward signals. In *2010 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4822–4829, May 2010.

[77] Eduardo Izquierdo and Thomas Bührmann. Analysis of a Dynamical Recurrent Neural Network Evolved for Two Qualitatively Different Tasks: Walking and Chemotaxis. In *ALIFE*, pages 257–264, 2008.

[78] Robert R. Jackson and R. Stimson Wilcox. Spider Flexibly Chooses Aggressive Mimicry Signals for Different Prey By Trial and Error. *Behaviour*, 127(1):21–36, January 1993.

[79] Nick Jakobi. Evolutionary Robotics and the Radical Envelope-of-Noise Hypothesis. *Adaptive Behavior*, 6(2):325–368, September 1997.

[80] F. J. Janzen and P. C. Phillips. Exploring the evolution of environmental sex determination, especially in reptiles. *Journal of Evolutionary Biology*, 19(6):1775–1784, November 2006.

[81] Thomas E. Jerde, John F. Soechting, and Martha Flanders. Coarticulation in fluent fingerspelling. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 23(6):2383–2393, March 2003.

[82] Jun Tani and M. Ito. Self-organization of behavioral primitives as multiple attractor dynamics: A robot experiment. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 33(4):481–488, 2003.

[83] Fred Keijzer. Representation in dynamical and embodied cognition. *Cognitive Systems Research*, 3(3):275–288, September 2002.

[84] Marc W. Kirschner and John C. Gerhart. *The Plausibility of Life: Resolving Darwin's Dilemma*. Yale University Press, New Haven, cincias biol-gicas e edition edition, November 2006.

[85] Jens Kober, J. Andrew Bagnell, and Jan Peters. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, 32(11):1238–1274, September 2013.

[86] Etienne Koechlin, Chrystèle Ody, and Frédérique Kouneiher. The Architecture of Cognitive Control in the Human Prefrontal Cortex. *Science*, 302(5648):1181–1185, November 2003.

[87] Petr E. Komers. Behavioural plasticity in variable environments. *Canadian Journal of Zoology*, 75(2):161–169, February 1997.

[88] George Konidaris, Scott Kuindersma, Roderic Grupen, and Andre S. Barreto. Constructing skill trees for reinforcement learning agents from demonstration trajectories. In *Advances in neural information processing systems*, pages 1162–1170, 2010.

[89] George Konidaris, Scott Kuindersma, Roderic Grupen, and Andrew Barto. Robot Learning from Demonstration by Constructing Skill Trees. *Int. J. Rob. Res.*, 31(3):360–375, March 2012.

[90] Sylvain Koos, Jean-Baptiste Mouret, and Stéphane Doncieux. Crossing the Reality Gap in Evolutionary Robotics by Promoting Transferable Controllers. In

Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation, GECCO '10, pages 119–126, New York, NY, USA, 2010. ACM.

[91] P. Kormushev, B. Ugurlu, S. Calinon, N.G. Tsagarakis, and D.G. Caldwell. Bipedal walking energy minimization by reinforcement learning with evolving policy parameterization. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 318–324, September 2011.

[92] Petar Kormushev, Sylvain Calinon, and Darwin Caldwell. Reinforcement Learning in Robotics: Applications and Real-World Challenges. *Robotics*, 2(3):122–148, July 2013.

[93] Wei-Po Lee, J. Hallam, and H.H. Lund. Applying genetic programming to evolve behavior primitives and arbitrators for mobile robots. In *, IEEE International Conference on Evolutionary Computation, 1997*, pages 501–506, April 1997.

[94] Joel Lehman and Risto Miikkulainen. Neuroevolution. *Scholarpedia*, 8(6):30977, 2013.

[95] Joel Lehman and Risto Miikkulainen. Overcoming Deception in Evolution of Cognitive Behaviors. In *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, GECCO '14, pages 185–192, New York, NY, USA, 2014. ACM.

[96] Joel Lehman and Kenneth O. Stanley. Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary computation*, 19(2):189–223, 2011.

[97] Dan Lessin, Don Fussell, and Risto Miikkulainen. Open-ended Behavioral Complexity for Evolved Virtual Creatures. In *Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation*, GECCO '13, pages 335–342, New York, NY, USA, 2013. ACM.

[98] Yun Lin, Shaogang Ren, M. Clevenger, and Yu Sun. Learning grasping force

from demonstration. In *2012 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1526–1531, May 2012.

[99] Fredrik Lin\a aker and Henrik Jacobsson. Mobile Robot Learning of Delayed Response Tasks Through Event Extraction: A Solution to the Road Sign Problem and Beyond. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'01, pages 777–782, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

[100] Jochen Maaß, Nnamdi Kohn, and Jürgen Hesselbach. Open Modular Robot Control Architecture for Assembly Using the Task Frame Formalism. *International Journal of Advanced Robotic Systems*, 3(1):2, March 2006.

[101] Naoki Maeda. External working memory and the amount of distributed cognition. In *Proceedings of the 34th Annual Meeting of the Cognitive Science Society*, pages 1954–1959, 2012.

[102] Andrea Maesani, Pradeep Ruben Fernando, and Dario Floreano. Artificial Evolution by Viability Rather than Competition. *PLOS ONE*, 9(1):e86831, January 2014.

[103] S. Magnenat, P. Rétornaz, M. Bonani, V. Longchamp, and F. Mondada. ASEBA: A Modular Architecture for Event-Based Control of Complex Robots. *IEEE/ASME Transactions on Mechatronics*, 16(2):321–329, April 2011.

[104] Laura Martin, Bulcsú Sándor, and Claudius Gros. Closed-loop Robots Driven by Short-Term Synaptic Plasticity: Emergent Explorative vs. Limit-Cycle Locomotion. *Frontiers in Neurorobotics*, 10, 2016.

[105] Georg Martius, Ralf Der, and J. Michael Herrmann. Robot Learning by Guided Self-Organization. In Mikhail Prokopenko, editor, *Guided Self-Organization: Inception*, number 9 in Emergence, Complexity and Computation, pages 223–260. Springer Berlin Heidelberg, 2014. DOI: 10.1007/978-3-642-53734-9_8.

[106] Gianluca Massera, Tomassino Ferrauto, Onofrio Gigliotta, and Stefano Nolfi. FARSA: An Open Software Tool for Embodied Cognitive Science. In *Advances in Artificial Life, ECAL 2013*, pages 538–545. MIT Press, September 2013.

[107] Gianluca Massera, Tomassino Ferrauto, Onofrio Gigliotta, and Stefano Nolfi. Designing adaptive humanoid robots through the FARSA open-source framework. *Adaptive Behavior*, 22(4):255–265, August 2014.

[108] M. K. McBeath, D. M. Shaffer, and M. K. Kaiser. How baseball outfielders determine where to run to catch fly balls. *Science (New York, N.Y.)*, 268(5210):569–573, April 1995.

[109] Jean-Arcady Meyer. Artificial Life and the Animat Approach to Artificial Intelligence. In *Artificial Intelligence*, pages 325–354. Elsevier, 1996. DOI: 10.1016/B978-012161964-0/50013-3.

[110] Orazio Miglino, Henrik Hautop Lund, and Stefano Nolfi. Evolving Mobile Robots in Simulated and Real Environments. *Artificial Life*, 2(4):417–434, July 1995.

[111] S. D. Mitchell. The units of behaviour in evolutionary explanations. In *In: Bekoff M, Jamieson D (eds) , Readings in Animal Cognition*. MIT Press, Cambridge, Mass, 1996.

[112] Jared M. Moore and Adam Stanton. Lexicase selection outperforms previous strategies for incremental evolution of virtual creature controllers. In *Advances in Artificial Life, ECAL 2017: Proceedings of the Fourteenth European Conference on the Synthesis and Simulation of Living Systems*, pages 290–297. MIT Press, September 2017.

[113] Michael E. Moran. The da Vinci Robot. *Journal of Endourology*, 20(12):986–990, December 2006.

[114] G. E. Nasr, E. A. Badr, and C. Joun. Cross Entropy Error Function in Neural Networks: Forecasting Gasoline Demand. In *Proceedings of the Fifteenth Inter-*

national Florida Artificial Intelligence Research Society Conference, pages 381–384. AAAI Press, 2002.

[115] Joseph Needham. *Science and Civilisation in China, Vol. 2, History of Scientific Thought*. Cambridge University Press, 1956.

[116] Andrew L. Nelson, Gregory J. Barlow, and Lefteris Doitsidis. Fitness functions in evolutionary robotics: A survey and analysis. *Robotics and Autonomous Systems*, 57(4):345–370, April 2009.

[117] Stefano Nolfi. Evolving robots able to self-localize in the environment: the importance of viewing cognition as the result of processes occurring at different time-scales. *Connection Science*, 14(3):231–244, 2002.

[118] Stefano Nolfi. Power and the limits of reactive agents. *Neurocomputing*, 42(1–4):119–145, January 2002.

[119] Stefano Nolfi. Behavior and Cognition as a Complex Adaptive System: Insights from Robotic Experiments. In *Philosophy of Complex Systems*, volume 10 of *Handbook of the Philosophy of Science*, pages 443–463. North-Holland, Amsterdam, 2011.

[120] Stefano Nolfi, Josh Bongard, Phil Husbands, and Dario Floreano. Evolutionary Robotics. In Bruno Siciliano and Oussama Khatib, editors, *Springer Handbook of Robotics*, pages 2035–2068. Springer International Publishing, 2016. DOI: 10.1007/978-3-319-32552-1_76.

[121] Stefano Nolfi and Dario Floreano. How co-evolution can enhance the adaptive power of artificial evolution: Implications for evolutionary robotics. In *Evolutionary Robotics*, pages 22–38. Springer, Berlin, Heidelberg, April 1998.

[122] Stefano Nolfi and Dario Floreano. Learning and Evolution. *Autonomous Robots*, 7(1):89–113, 1999.

[123] Stefano Nolfi and Dario Floreano. *Evolutionary Robotics: The Biology,Intelligence,and Technology*. MIT Press, Cambridge, MA, USA, 2000.

[124] Stefano Nolfi and Davide Marocco. Active perception: A sensorimotor account of object categorization. In *From animals to animats 7: Proceedings of the VII international conference on simulation of adaptive behavior*, pages 266–271. et al., Cambridge, MA, 2002.

[125] Stefano Nolfi and Domenico Parisi. Learning to Adapt to Changing Environments in Evolving Neural Networks. In *Adaptive Behavior*, pages 75–98, 1997.

[126] Nikita Noskov, Evert Haasdijk, Berend Weel, and A. E. Eiben. MONEE: Using Parental Investment to Combine Open-Ended and Task-Driven Evolution. In *Applications of Evolutionary Computation*, Lecture Notes in Computer Science, pages 569–578. Springer, Berlin, Heidelberg, April 2013.

[127] Charles Ollion, Tony Pinville, and Stéphane Doncieux. With a little help from selection pressures: evolution of memory in robot controllers. In *Artificial Life*, volume 13, pages 407–414. MIT Press, July 2012.

[128] A. M. M. Omer, R. Ghorbani, Hun-ok Lim, and Atsuo Takanishi. Semi-passive dynamic walking for biped walking robot using controllable joint stiffness based on dynamic simulation. In *2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pages 1600–1605, July 2009.

[129] Daniel Otte. Simple Versus Elaborate Behavior in Grasshoppers an Analysis of Communication in the Genus Syrbula. *Behaviour*, 42(3):291–321, January 1972.

[130] P. Y. Oudeyer, F. Kaplan, and V. V. Hafner. Intrinsic Motivation Systems for Autonomous Mental Development. *IEEE Transactions on Evolutionary Computation*, 11(2):265–286, April 2007.

[131] Rainer W. Paine and Jun Tani. How Hierarchical Control Self-organizes in Artificial Adaptive Systems. *Adaptive Behavior*, 13(3):211–225, September 2005.

[132] T. Palleja, M. Tresanchez, M. Teixido, and J. Palacin. Modeling floor-cleaning coverage performances of some domestic mobile robots in a reduced scenario. *Robotics and Autonomous Systems*, 58(1):37–45, January 2010.

[133] Ezequiel Di Paolo, Thomas Buhrmann, and Xabier Barandiaran. *Sensorimotor Life: An Enactive Proposal*. Oxford University Press, May 2017. Google-Books-ID: ogOpDgAAQBAJ.

[134] Giancarlo Petrosino, Domenico Parisi, and Stefano Nolfi. Selective attention enables action selection: evidence from evolutionary robotics experiments. *Adaptive Behavior*, 21(5):356–370, October 2013.

[135] Rolf Pfeifer and Josh Bongard. *How the Body Shapes the Way We Think: A New View of Intelligence*. MIT Press, October 2006.

[136] Brad E Pfeiffer and David J Foster. Hippocampal place cell sequences depict future paths to remembered goals. *Nature*, 497(7447):74–79, May 2013.

[137] Robert F. Port and Timothy Van Gelder. *Mind as Motion: Explorations in the Dynamics of Cognition*. MIT Press, Cambridge, MA, 1995.

[138] Mitchell A. Potter and Kenneth A. De Jong. Cooperative Coevolution: An Architecture for Evolving Coadapted Subcomponents. *Evolutionary Computation*, 8(1):1–29, March 2000.

[139] Tony Prescott. Action selection. *Scholarpedia*, 3(2):2705, 2008.

[140] C. J. Pretorius, M. C. du Plessis, and C. B. Cilliers. Simulating Robots Without Conventional Physics: A Neural Network Approach. *Journal of Intelligent & Robotic Systems*, 71(3-4):319–348, September 2013.

[141] Y. Qu, S. Pandhiti, K. S. Bullard, W. D. Potter, and K. F. Fezer. Development of a Genetic Fuzzy Controller for an Unmanned Aerial Vehicle. In Kishan G. Mehrotra, Chilukuri K. Mohan, Jae C. Oh, Pramod K. Varshney, and Moonis

Ali, editors, *Modern Approaches in Applied Intelligence*, number 6704 in Lecture Notes in Computer Science, pages 328–335. Springer Berlin Heidelberg, 2011.

[142] Saufiah A. Rahim, Azman M. Yusof, and Thomas Bräunl. Genetically Evolved Action Selection Mechanism in a Behavior-based System for Target Tracking. *Neurocomput.*, 133:84–94, June 2014.

[143] S. D. Rajan. Sizing, Shape, and Topology Design Optimization of Trusses Using Genetic Algorithm. *Journal of Structural Engineering*, 121(10):1480–1487, October 1995.

[144] Evan F. Risko and Timothy L. Dunn. Storing information in-the-world: Metacognition and cognitive offloading in a short-term memory task. *Consciousness and Cognition*, 36:61–74, November 2015.

[145] Evan F. Risko and Sam J. Gilbert. Cognitive Offloading. *Trends in Cognitive Sciences*, 20(9):676–688, September 2016.

[146] Evan F. Risko, Srdan Medimorec, Joseph Chisholm, and Alan Kingstone. Rotating With Rotated Text: A Natural Behavior Approach to Investigating Cognitive Offloading. *Cognitive Science*, 38(3):537–564, April 2014.

[147] R. M. Rylatt and C. A. Czarnecki. Embedding Connectionist Autonomous Agents in Time: The 'Road Sign Problem'. *Neural Processing Letters*, 12(2):145–158, 2000.

[148] Khairulmizam Samsudin, Faisul Arif Ahmad, and Syamsiah Mashohor. A highly interpretable fuzzy rule base using ordinal structure for obstacle avoidance of mobile robot. *Applied Soft Computing*, 11(2):1631–1637, March 2011.

[149] Kirk Y. W. Scheper, Sjoerd Tijmons, Cornelis C. de Visser, and Guido C. H. E. de Croon. Behavior Trees for Evolutionary Robotics. *Artificial Life*, 22(1):23–48, November 2015.

[150] Jürgen Schmidhuber. A Possibility for Implementing Curiosity and Boredom in Model-building Neural Controllers. In *Proceedings of the First International Conference on Simulation of Adaptive Behavior on From Animals to Animats*, pages 222–227, Cambridge, MA, USA, 1990. MIT Press.

[151] J. Schrum and R. Miikkulainen. Evolving Multimodal Networks for Multitask Games. *IEEE Transactions on Computational Intelligence and AI in Games*, 4(2):94–111, June 2012.

[152] Anil Seth, Tony Prescott, and Joanna Bryson, editors. *Modelling Natural Action Selection*. Cambridge University Press, Cambridge, 2011.

[153] Anil K. Seth. Optimised agent-based modelling of action selection. In *Modelling Natural Action Selection*. Cambridge University Press, 2011.

[154] Arthur M. Shapiro. Seasonal Polyphenism. In Max K. Hecht, William C. Steere, and Bruce Wallace, editors, *Evolutionary Biology*, pages 259–333. Springer US, 1976. DOI: 10.1007/978-1-4615-6950-3_6.

[155] Andrew C. Slocum, Douglas C. Downey, Randall D. Beer, and All D. Beer. Further Experiments in the Evolution of Minimally Cognitive Behavior: From Perceiving Affordances to Selective Attention. In *From animals to animats 6: Proceedings of the sixth international conference on simulation of adaptive behavior*, pages 430–439. MIT Press, 2000.

[156] Valerio Sperati, Vito Trianni, and Stefano Nolfi. Evolving coordinated group behaviours through maximisation of mean mutual information. *Swarm Intelligence*, 2(2-4):73–95, September 2008.

[157] S. W. Squyres, A. H. Knoll, R. E. Arvidson, J. W. Ashley, J. F. Bell, W. M. Calvin, P. R. Christensen, B. C. Clark, B. A. Cohen, P. A. de Souza, L. Edgar, W. H. Farrand, I. Fleischer, R. Gellert, M. P. Golombek, J. Grant, J. Grotzinger, A. Hayes, K. E. Herkenhoff, J. R. Johnson, B. Jolliff, G. Klingelhöfer, A. Knudson, R. Li,

T. J. McCoy, S. M. McLennan, D. W. Ming, D. W. Mittlefehldt, R. V. Morris, J. W. Rice, C. Schröder, R. J. Sullivan, A. Yen, and R. A. Yingst. Exploration of Victoria Crater by the Mars Rover Opportunity. *Science*, 324(5930):1058–1061, May 2009.

[158] Russell K. Standish. Open-ended artificial evolution. *International Journal of Computational Intelligence and Applications*, 3(02):167–175, 2003.

[159] Kenneth O. Stanley. Compositional Pattern Producing Networks: A Novel Abstraction of Development. *Genetic Programming and Evolvable Machines*, 8(2):131–162, June 2007.

[160] Kenneth O. Stanley and Risto Miikkulainen. Evolving neural networks through augmenting topologies. *Evol. Comput.*, 10(2):99–127, June 2002.

[161] Kenneth O. Stanley and Risto Miikkulainen. Competitive Coevolution Through Evolutionary Complexification. *Journal of Artificial Intelligence Research*, 21(1):63–100, February 2004.

[162] Adam Stanton and A. D. Channon. Incremental neuroevolution of reactive and deliberative 3d agents. In *Advances in Artificial Life, ECAL 2015: Proceedings of the Thirteenth European Conference on the Synthesis and Simulation of Living Systems*, pages 341–348. Keele University, 2015.

[163] Adam Stanton and Alastair Channon. Heterogeneous complexification strategies robustly outperform homogeneous strategies for incremental evolution. In *Advances in Artificial Life, ECAL 2013: Proceedings of the Twelfth European Conference on the Synthesis and Simulation of Living Systems*, volume 12, pages 973–980, September 2013.

[164] Silke Steingrube, Marc Timme, Florentin Wörgötter, and Poramate Manoonpong. Self-organized adaptation of a simple neural circuit enables complex robot behaviour. *Nature Physics*, 6(3):224–230, March 2010.

[165] John Robert Stewart, Olivier Gapenne, and Ezequiel A. Di Paolo. *Enaction: Toward a New Paradigm for Cognitive Science*. MIT Press, Cambridge, MA, 2010.

[166] Peter Stone and Manuela Veloso. Layered Learning. In Ramon López de Mántaras and Enric Plaza, editors, *Machine Learning: ECML 2000*, number 1810 in Lecture Notes in Computer Science, pages 369–381. Springer Berlin Heidelberg, May 2000. DOI: 10.1007/3-540-45164-1_38.

[167] Y. Sugimoto and K. Osuka. Walking Control of Quasi Passive Dynamic Walking Robot "Quartet III" based on Continuous Delayed Feedback Control. In *2004 IEEE International Conference on Robotics and Biomimetics*, pages 606–611, August 2004.

[168] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: an introduction*. Adaptive computation and machine learning. MIT Press, Cambridge, Mass, 1998.

[169] J. Tani and S. Nolfi. Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems. *Neural Networks*, 12(7–8):1131–1141, October 1999.

[170] Esther Thelen and Linda B. Smith. *A Dynamic Systems Approach to the Development of Cognition and Action*. MIT Press, Cambridge, MA, January 1996.

[171] Serge Thill, Daniele Caligiore, Anna M. Borghi, Tom Ziemke, and Gianluca Baldassarre. Theories and computational models of affordance and mirror systems: An integrative review. *Neuroscience & Biobehavioral Reviews*, 37(3):491–521, March 2013.

[172] Julian Togelius. Evolution of a subsumption architecture neurocontroller. *Journal of Intelligent and Fuzzy Systems: Applications in Engineering and Technology*, 15(1):15–20, January 2004.

[173] Elio Tuci, Matt Quinn, and Inman Harvey. An Evolutionary Ecological Approach to the Study of Learning Behavior Using a Robot-Based Model. *Adaptive Behavior*, 10(3-4):201–221, July 2002.

[174] Barbara Tversky. Visualizing Thought. *Topics in Cognitive Science*, 3(3):499–535, July 2011.

[175] Joseba Urzelai and Dario Floreano. Evolutionary Robotics: Coping with Environmental Change. In *Proceedings of the 2Nd Annual Conference on Genetic and Evolutionary Computation*, GECCO'00, pages 941–948, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers Inc.

[176] A. Vakanski, I. Mantegh, A. Irish, and F. Janabi-Sharifi. Trajectory Learning for Robot Programming by Demonstration Using Hidden Markov Model and Dynamic Time Warping. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 42(4):1039–1052, August 2012.

[177] Francisco J. Varela, Eleanor Rosch, and Evan Thompson. *The Embodied Mind: Cognitive Science and Human Experience*. Mit Press, Cambridge, MA, February 1993.

[178] Vassilis Vassiliades and Chris Christodoulou. Behavioral plasticity through the modulation of switch neurons. *Neural Networks*, 74:35–51, February 2016.

[179] Roby Velez and Jeff Clune. Novelty Search Creates Robots with General Skills for Exploration. In *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*, GECCO '14, pages 737–744, New York, NY, USA, 2014. ACM.

[180] Phillip Verbancsics and Kenneth O. Stanley. Constraining Connectivity to Encourage Modularity in HyperNEAT. In *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, GECCO '11, pages 1483–1490, New York, NY, USA, 2011. ACM.

[181] C. von Hofsten and L. Rönnqvist. Preparation for grasping an object: a developmental study. *Journal of Experimental Psychology. Human Perception and Performance*, 14(4):610–621, November 1988.

[182] Michael L. Wehmeyer. *The Oxford Handbook of Positive Psychology and Disability*. OUP USA, September 2013. Google-Books-ID: ESyRAAAAQBAJ.

[183] R. Wehner. Desert ant navigation: how miniature brains solve complex tasks. *Journal of Comparative Physiology A*, 189(8):579–588, August 2003.

[184] John W. Wenzel. Application of the biogenetic law to behavioral ontogeny: a test using nest architecture in paper wasps. *Journal of Evolutionary Biology*, 6(2):229–247, March 1993.

[185] Mary Jane West-Eberhard. *Developmental Plasticity and Evolution*. OUP USA, Oxford ; New York, February 2003.

[186] Anne E. Wignall, Robert R. Jackson, R. Stimson Wilcox, and Phillip W. Taylor. Exploitation of environmental noise by an araneophagic assassin bug. *Animal Behaviour*, 82(5):1037–1042, November 2011.

[187] Andrew M Wikenheiser and A David Redish. Decoding the cognitive map: ensemble hippocampal sequences and decision making. *Current Opinion in Neurobiology*, 32:8–15, June 2015.

[188] George Christopher Williams. *Adaptation and natural selection: a critique of some current evolutionary thought*. Princeton science library. Princeton Univ. Press, Princeton, NJ, 1996. OCLC: 833082108.

[189] Paul Williams and Randall Beer. Environmental feedback drives multiple behaviors from the same neural circuit. In *Advances in Artificial Life, ECAL*, volume 12, pages 268–275, 2013.

[190] Margaret Wilson. Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4):625–636, 2002.

[191] Emma R. Wood, Paul A. Dudchenko, R. Jonathan Robitsek, and Howard Eichenbaum. Hippocampal Neurons Encode Information about Different Types of Memory Episodes Occurring in the Same Location. *Neuron*, 27(3):623–633, September 2000.

[192] Yoshiyuki Yamamoto and William R. Jeffery. Central Role for the Lens in Cave Fish Eye Degeneration. *Science*, 289(5479):631–633, July 2000.

[193] Yoshiyuki Yamamoto, David W. Stock, and William R. Jeffery. Hedgehog signalling controls eye degeneration in blind cavefish. *Nature*, 431(7010):844–847, October 2004.

[194] Philipp Zech, Simon Haller, Safoura Rezapour Lakani, Barry Ridge, Emre Ugur, and Justus Piater. Computational models of affordance in robotics: a taxonomy and systematic classification. *Adaptive Behavior*, 25(5):235–271, October 2017.

[195] Tom Ziemke and Mikael Thieme. Neuromodulation of Reactive Sensorimotor Mappings as a Short-Term Memory Mechanism in Delayed Response Tasks. *Adaptive Behavior*, 10(3-4):185–199, July 2002.

Bound copies of published papers

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 $\times$ 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 $\pm$33 and $\pm$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

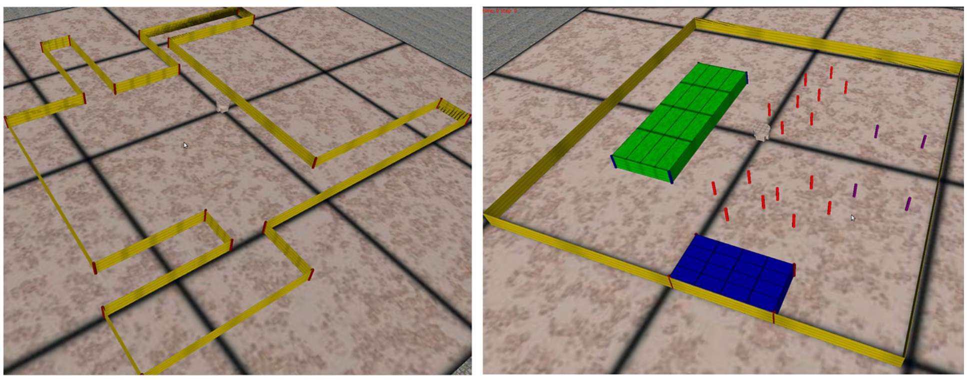


Fig. 1 Examples of concave and convex environments, *left* and *right*, respectively

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.

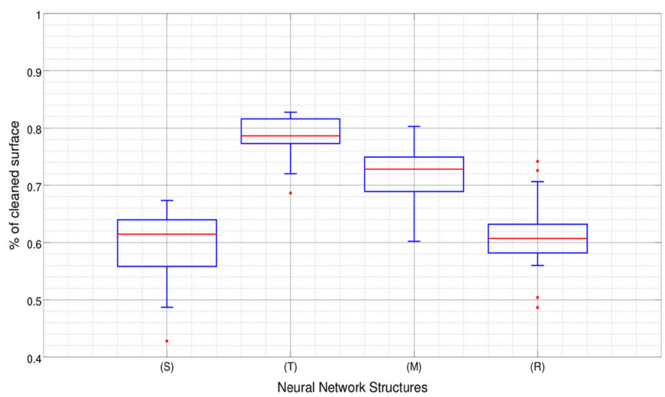
In “On the mechanisms supporting behaviour differentiation and arbitration”, we will discuss the mechanisms that support behavioural differentiation and arbitration by analysing the behavioural solutions found in the different experimental conditions. As we will see, the two most important mechanisms 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



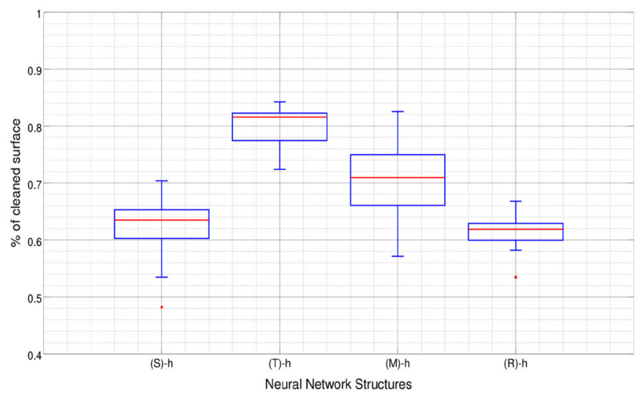


Fig. 2 *Box plots* of performance in the concave environment. The *left* and *right* figures report the results obtained without internal neurons and with internal neurons, respectively. The *box plots* display the performance of the best robot of the last generation in the four experimental conditions, 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

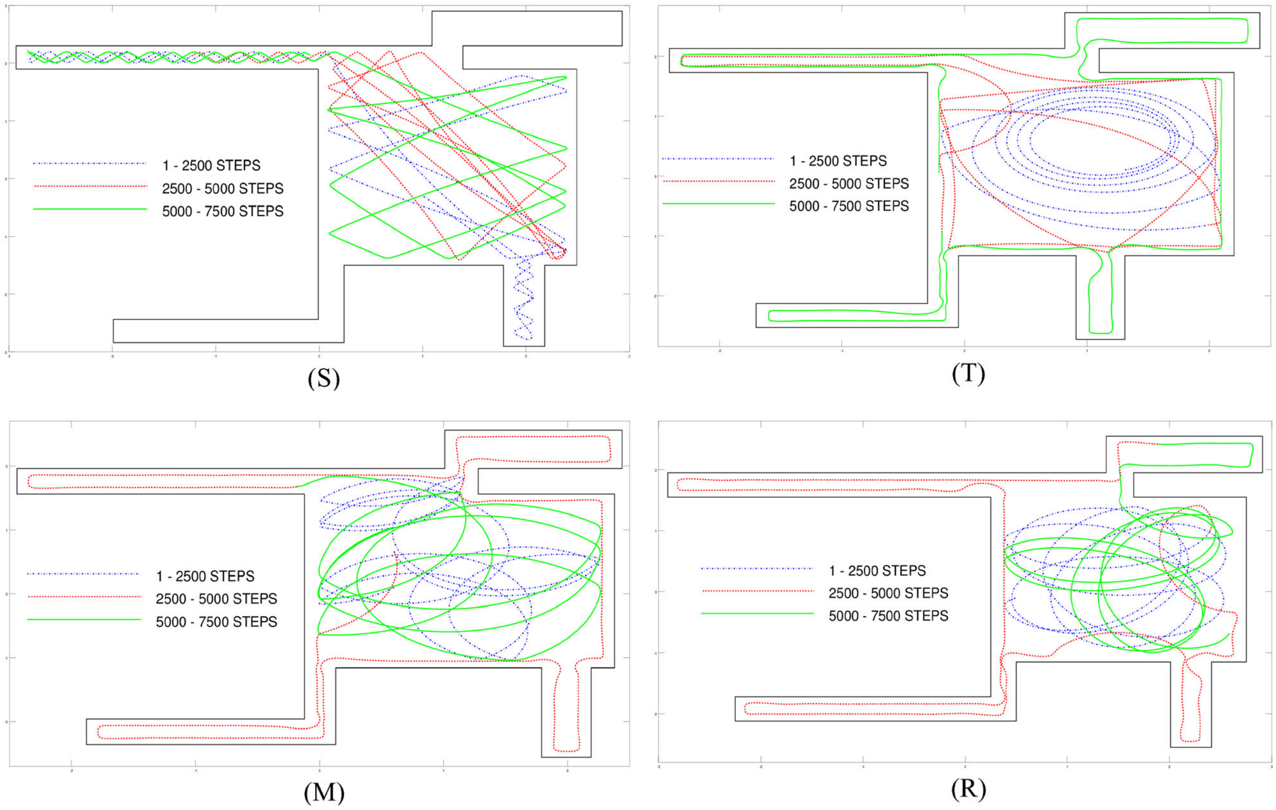


Fig. 3 Typical trajectories displayed by the best robots of the four experimental conditions without hidden units in the concave environment. The portions of the trajectory produced during the first, second, and third part of the trial (i.e., from step 1 to 2500, from step 2501 to 5000, and from step 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 ± 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

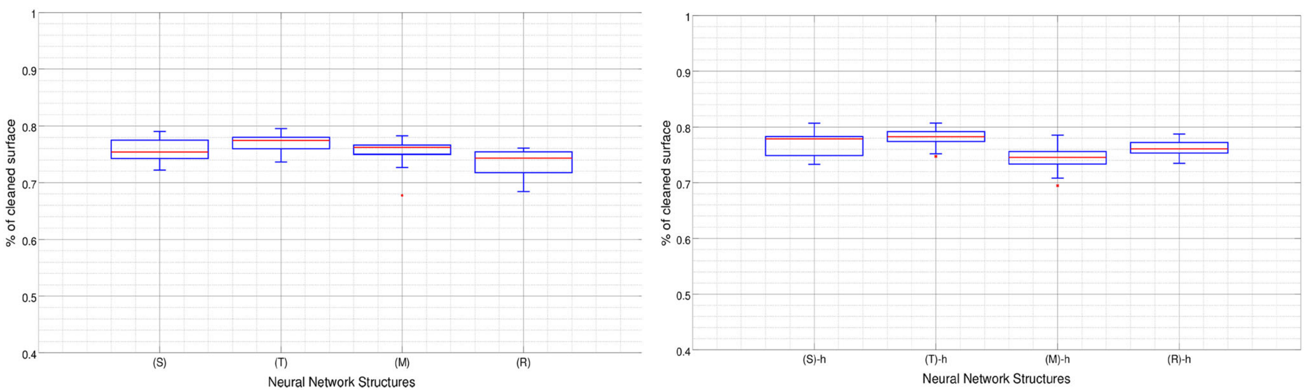


Fig. 4 *Box plots* of performance in the convex environment. The *left* and *right* figures report the results obtained without internal neurons and with internal neurons, respectively

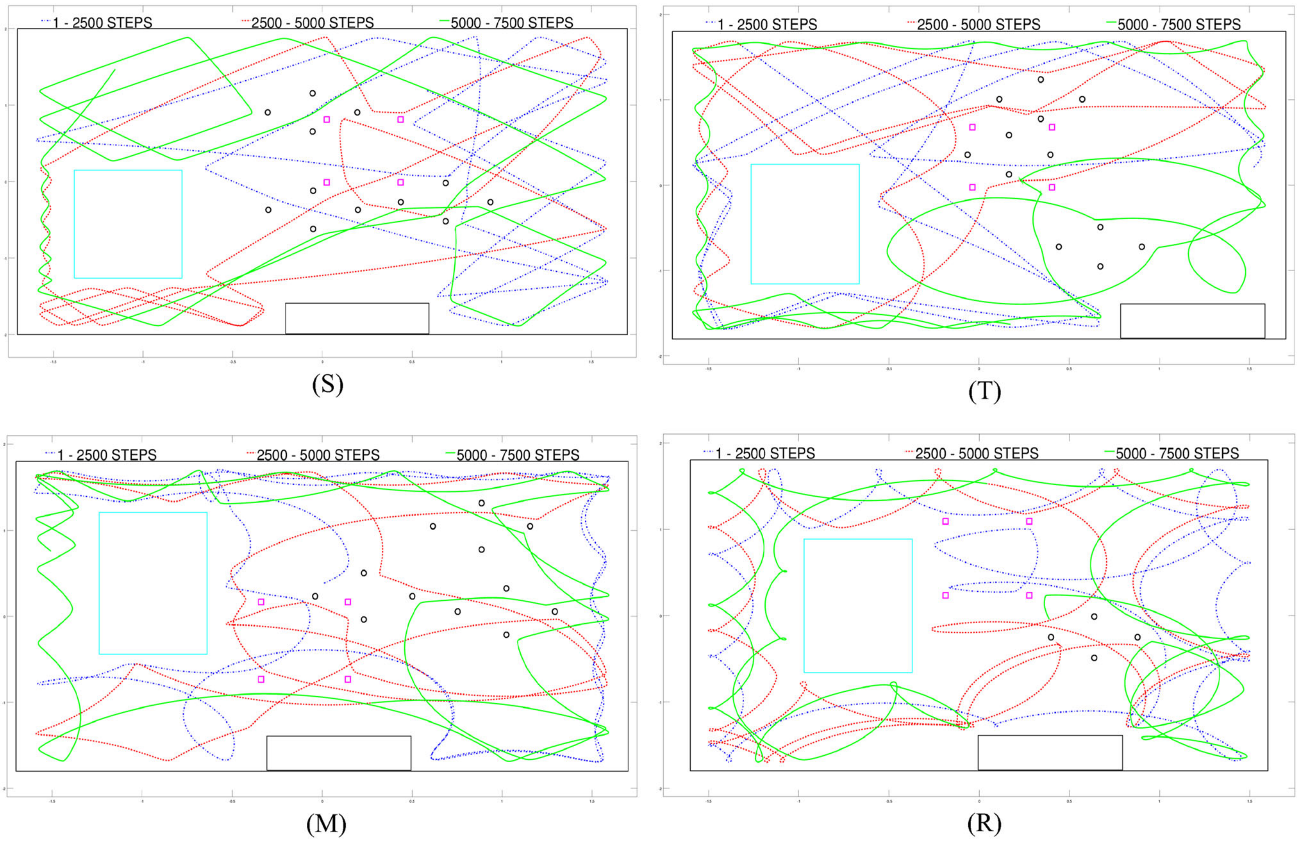


Fig. 5 Typical trajectory displayed by the best robots of the four experimental conditions without hidden units in the convex environment

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

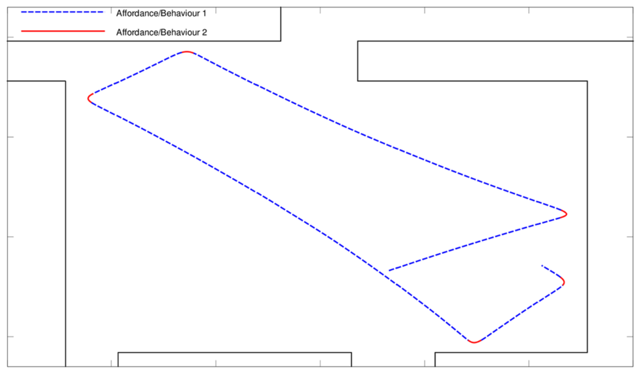


Fig. 6 Exemplification of short-term behavioural plasticity in the case of an exploration behaviour that is realised by alternating a move-forward and an obstacle-avoidance behaviour (shown in *blue* and *red*, respectively). The former behaviour is elicited by perceptual states in which the frontal 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

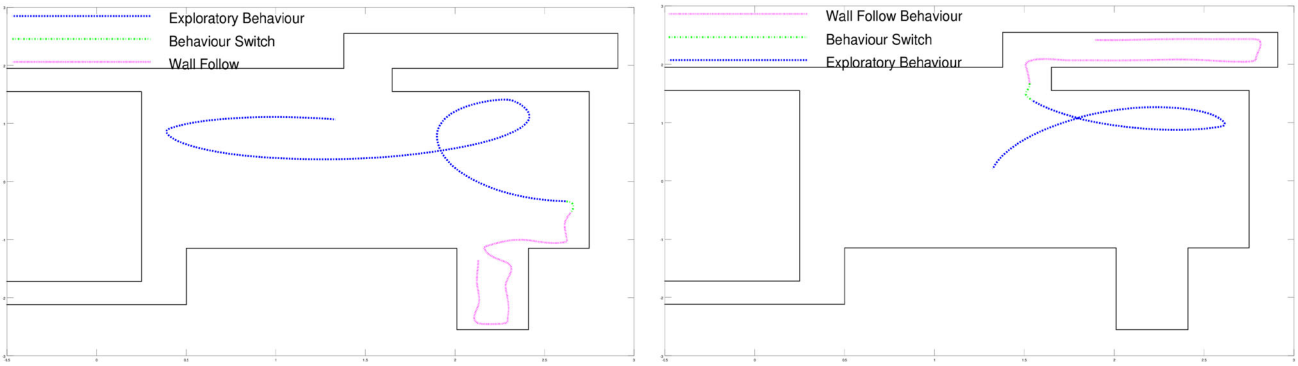


Fig. 7 Illustration of how the best (R) robot switches from the exploration to the wall-following behaviour and vice versa (*left* and *right*, respectively)

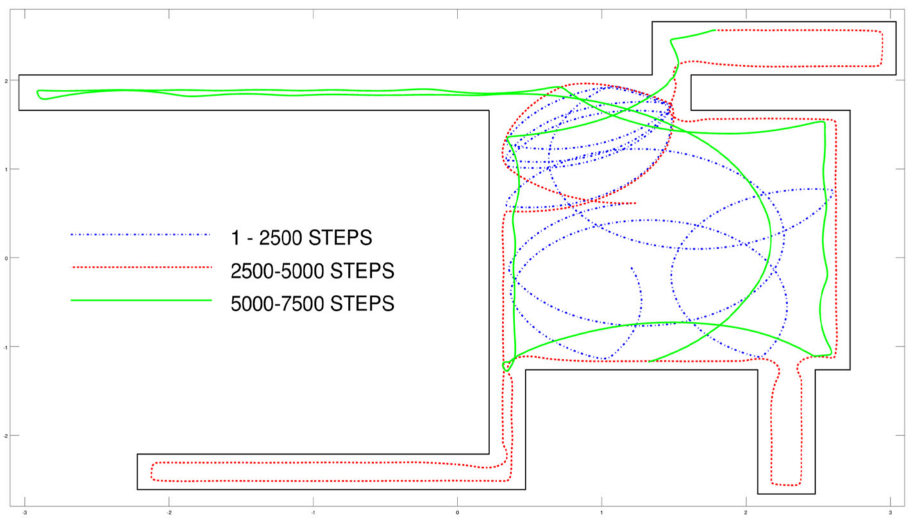


Fig. 8 Trajectory produced by the best (M) robot which produces an exploration behaviour under the control of the first neural module, an exploration and then a wall-following behaviour under the control of the second neural module, and a wall-following and then an exploration behaviour under the control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

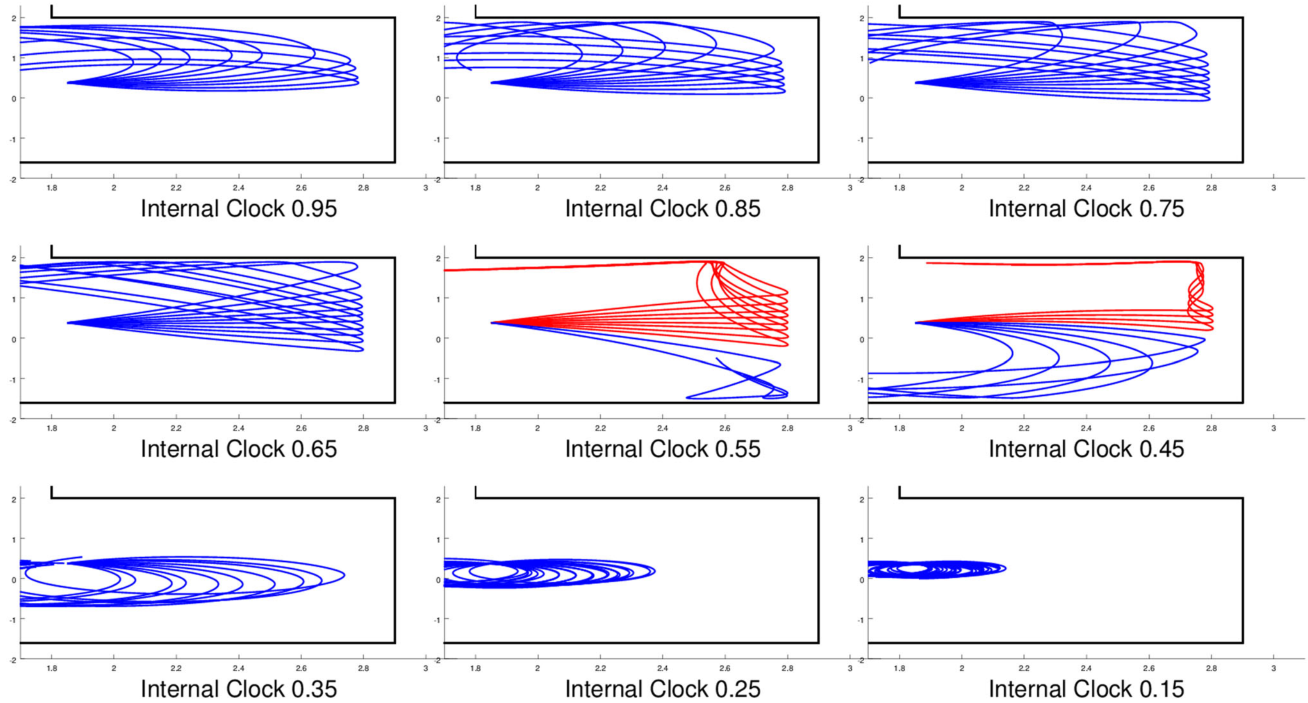


Fig. 9 Behaviour produced by the best (T) robot during different trials in which it started from the same initial position with systematically varied orientations and systematically varied state of the time neuron. The *red* and *blue lines* represent the trajectories produced by the robot during trials 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

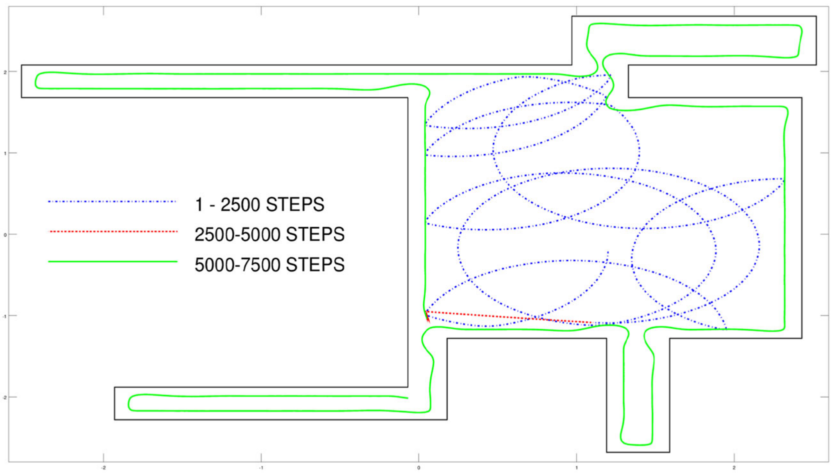


Fig. 10 Trajectory produced by one of the three best (M) robots characterised by a second module that is specialised for enabling a suitable transition from the exploration to the wall-following behaviour

from which the wall-following behaviour can be effectively triggered during the critical period (i.e., during 3.17 ± 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

RESEARCH ARTICLE

Cognitive Offloading Does Not Prevent but Rather Promotes Cognitive Development

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

1 Institute of Cognitive Sciences and Technologies, National Research Council (CNR), Via S. Martino della Battaglia, 44, 00185, Roma, Italia, 2 Center for Computational Sciences (C3), Federal University of Rio Grande (FURG), Av. Italia, km 8, 96203–900, Rio Grande, Brasil

* jonatatyska@gmail.com (JTC); stefano.nolfi@istc.cnr.it (SN)

Citation: Carvalho JT, Nolfi S (2016) Cognitive Offloading Does Not Prevent but Rather Promotes Cognitive Development. PLoS ONE 11(8): e0160679. doi:10.1371/journal.pone.0160679

Editor: Stephen D Ginsberg, Nathan S Kline Institute, UNITED STATES

Received: March 14, 2016

Accepted: July 24, 2016

Published: August 9, 2016

Data Availability Statement: The experiments described in this paper can be replicated by downloading and installing FARSA and the associated experimental plugin from "https://sourceforge.net/projects/farsa/" and from "http://laral.istc.cnr.it/cognoffpone/dtmaze.tgz."

Funding: JTC has a PhD Scholarship provided by Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior (CAPES), Brazil (http://www.capes.gov.br), Grant number: 9052-13-5. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

We investigate the relation between the development of reactive and cognitive capabilities. In particular we investigate whether the development of reactive capabilities prevents or promotes the development of cognitive capabilities in a population of evolving robots that have to solve a time-delay navigation task in a double T-Maze environment. Analysis of the experiments reveals that the evolving robots always select reactive strategies that rely on cognitive offloading, i.e., the possibility of acting so as to encode onto the relation between the agent and the environment the states that can be used later to regulate the agent's behavior. The discovery of these strategies does not prevent, but rather facilitates, the development of cognitive strategies that also rely on the extraction and use of internal states. Detailed analysis of the results obtained in the different experimental conditions provides evidence that helps clarify why, contrary to expectations, reactive and cognitive strategies tend to have synergetic relationships.

Introduction

Developments in psychology, neuroscience, linguistics, robotics and philosophy have clarified that cognition cannot be studied properly without taking into sufficient account the role of the body, action and the external world [1–7]. The agent's body and the environment in which it is situated provide a great deal of structure that is used to operate appropriately. Consequently, in many cases the internal capabilities required are much simpler than those previously hypothesized within disembodied accounts. For example, moving around in a city does not necessarily require an elaborate representation of the city's layout. The ability to recognize a limited number of turning decision points combined with the ability to just follow the street between decision points might suffice [8]. Similarly, baseball players do not need to estimate the trajectory of the flying ball to be intercepted through complex calculations. They can simply adjust their running speed so as to maintain the relative angle between their eyes and the ball constant [9].

Exploitation of the information that can be extracted directly from the environment and of the effects of situated actions do not only affect the agents' low-level capabilities. Embodied and embedded strategies (like those described above) co-exist and interact with different strategies that are less dependent on agent/environmental interactions and more dependent on

Competing Interests: The authors have declared that no competing interests exist.

internal processes at all levels of organization [5]. However, the relation and the interaction between strategies and capabilities that differ in that respect have not yet been investigated. Consequently, the question of how these different types of strategies can be integrated from an operational and developmental perspective is still open. In particular, one important question that needs to be answered is the following: "Is cognition truly seamless–implying a gentle, incremental trajectory linking fully embodied responsiveness to abstract thought and off-line reason? Or is it a patchwork quilt, with jumps and discontinuities and with very different kinds of processing and representations serving different needs?" [5].

In this paper we investigate the relation between the development of reactive and cognitive capabilities. In particular we investigate whether the development of reactive capabilities prevents or promotes the development of cognitive capabilities. For the scope of this paper we define cognition as the ability to integrate sensory-motor information over time into internal states and to use these internal states to regulate the way the agent reacts to perceived stimuli. The term cognition is often used in a more restricted way. In the above definition, we focus on a fundamental capacity that is at the basis of all cognitive capabilities (e.g. perception, memory, attention, decision-making, reasoning, language, etc.).

Evolutionary Robotics [10] is a suitable method for studying the relation between reactive and cognitive abilities in adaptive agents. Indeed research carried out in this area has demonstrated how evolving robots can master both problems that have reactive solutions and problems that require the development of cognitive abilities (see for example [11–19]). However, what has not been sufficiently investigated to date is the relation between reactive and cognitive strategies.

Any adaptive problem typically admits qualitatively different sub-optimal and optimal solutions. Depending on the circumstances, the discovery of one type of sub-optimal solution might facilitate or block the discovery of better alternative solutions. Indeed, the discovery of sub-optimal solutions that cannot be progressively transformed into better solutions without loss of performance should retard or block the discovery of effective solutions. On the contrary, the selection of sub-optimal solutions that can be transformed into better solutions without causing significant performance loss can facilitate the discovery of effective solutions. In the latter case, the strength of the facilitation effect would depend on the level of overlap or similarity between the first and the second solutions.

Since reactive solutions are typically simpler than cognitive solutions from the point of view of the complexity of the required control mechanisms, and since adaptation tends to find the simpler solution first, the question is the following: Can reactive solutions be transformed into better solutions that also include the utilization of internal states without causing significant performance loss and what is the level of overlap/similarity between sub-optimal reactive solutions and better solutions that include cognitive capabilities.

According to some authors, the development of reactive solutions blocks the development of cognitive capabilities. In particular, [20] claimed that reactive solutions constitute hard to escape local optima that prevent the development of cognitive solutions. Similarly, [21] claimed that the deception caused by the availability of locally optimal reactive policies is one of the main factors that explains why it is difficult for cognitive policies to evolve. For these reasons the authors hypothesized that the development of cognitive solutions necessarily requires specific selective cognitive pressures such as fitness components that encourage the development of short-term memory [20] or mechanisms for avoiding local optima, such as novelty search [21].

One aspect that is particularly relevant from the viewpoint of the relation between reactive and cognitive strategies is cognitive offloading, i.e., the possibility of offloading cognitive work onto the environment [22–24]. In particular, the possibility of acting so as to encode the states that can be used to regulate the agent's behavior onto the external environment and/or onto the relation between the agent and the environment. In fact, the possibility of encoding the required states internally or externally suggests that cognitive strategies and reactive strategies (that rely on cognitive offloading) represent two alternative but functionally equivalent modalities. A simple example of cognitive offloading related to everyday human life is crossing two fingers so to avoid forgetting to perform a certain action [24–26]. An example of cognitive offloading realized in a robotic scenario consists of dropping markers in the environment that are used to find the way back to the home location [19].

Cognitive offloading is usually considered in the case of cognitive agents that already possess cognitive abilities but offload information into the environment or into their relation with the environment. In this work, instead, we focus on a developmental perspective in which the agents need to develop a certain skill to adapt to their environment and can do so by using a reactive strategy that relies on cognitive offloading, a cognitive strategy that relies on internal states, or on a hybrid strategy that relies on both.

In this paper we report a series of experiments in which a population of robots provided with neural network controllers were evolved for the ability to master a navigation problem in a double T-maze environment that required the exhibition of delayed response behavior. Analysis of the experiments reveals that the evolving robots always selected reactive strategies that relied on cognitive offloading. The discovery of these strategies did not prevent but rather facilitated the development of improved strategies that also relied on the extraction and use of internal states. A detailed analysis of the results obtained in different experimental conditions provides evidence that helps clarify why, contrary expectations, reactive and cognitive strategies tend to have synergetic relationships.

Method

A paradigmatic class of problems that require cognitive skills is constituted by delayed response tasks in which an agent has to act at a certain time t in a conditional dependent manner on the basis of stimuli it encountered at a previous time (t—delay). A simple example of a delayed response task is the so-called road sign task in which an agent that is initially located at the bottom of a T-Maze environment needs to travel toward the top-left or top-right destination by turning left or right at the T-junction depending on whether it previously experienced a stimulus on the left or on the right side of the central corridor, respectively (Fig 1). Therefore, this task was chosen by several researchers to study the evolution of cognitive robots [20, 27–31].

In reality, as demonstrated by [15] this problem has a non-cognitive solution, i.e., a reactive strategy in which the robot always acts on the basis of its current sensory state. Indeed, the robot could solve the task by approaching the experienced stimulus, when visible, and then by following the nearby left or right wall. The availability of simple and optimal reactive solutions of this type prevents the development of cognitive solutions. These reactive solutions, in fact, are easy to discover, optimal and, consequently, evolutionarily stable.

The question that remains open is whether cognitive solutions can evolve in experimental settings that do not allow for optimal reactive solutions or whether the discovery of sub-optimal reactive solutions prevents the discovery of better solutions [20–21] and consequently might drive the evolutionary process toward inescapable local optima. To investigate this question we decided to carry out the evolutionary experiments described in the following sub-sections.

The robot, the environment, and the task

The environment consisted of a double T-Maze (Fig 2) that included four different destinations and two types of stimuli that could be experienced in four different corresponding patterns

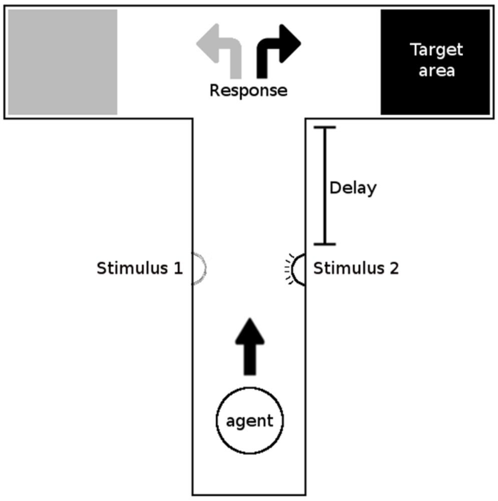


Fig 1. The T-Maze task. The bottom portion of the central corridor includes a stimulus located on the left or the right side. The position of the stimulus corresponds to the position of the target destination. Adapted from [29].

doi:10.1371/journal.pone.0160679.g001

(left-left, left-right, right-left, right-right). The increase in complexity with respect to the simple T-Maze environment was due to the fact that the number of destinations to be reached was higher, the number of stimuli experienced was higher, the number of stimuli-dependent decisions that had to be made was higher and the time delay between the moment in which the stimuli were experienced and the moment in which the stimuli-dependent decisions had to be made was longer.

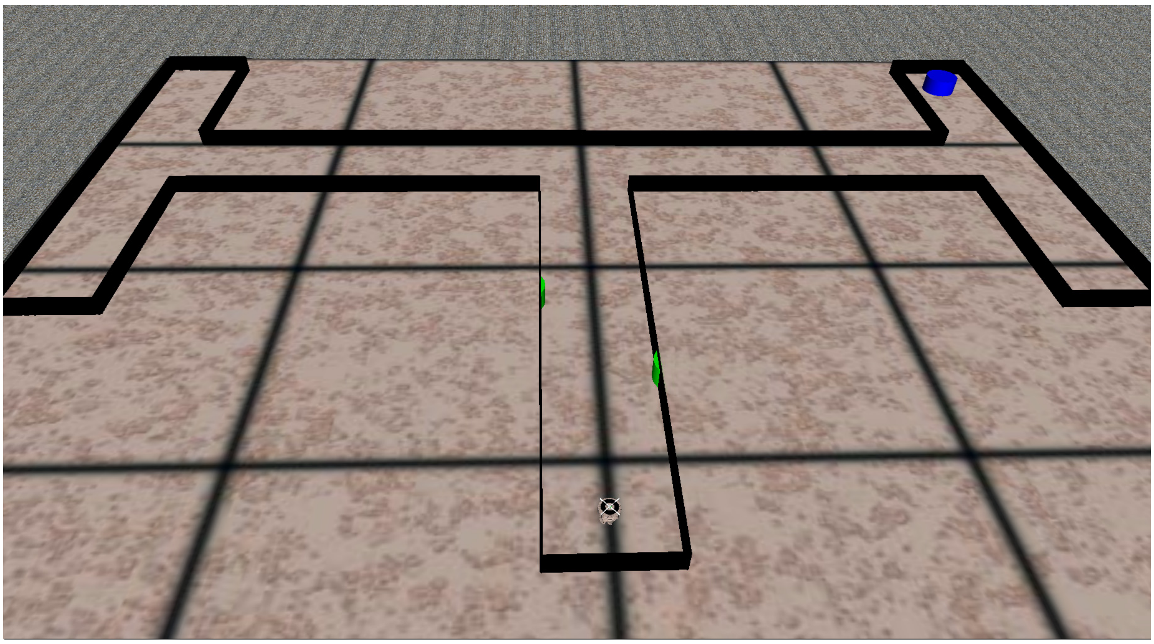


Fig 2. The double T-Maze task. The blue cylinder, which represented the target destination, was located at one of the four end points of the lateral corridors. The central corridor included two green stimuli located in the first and the second part of the corridor on the left or the right side. The position of the first and the second stimulus indicated whether the robot should turn left or right at the first and the second junction, respectively, to reach the target destination.

doi:10.1371/journal.pone.0160679.g002

To obtain solutions that were robust with respect to environmental variations, the initial positions and orientations of the robot and the size of the environment were randomly varied at the beginning of each trial. More precisely the length of the central corridor and the two vertical corridors was randomly set during each trial within 4.5m±0.5m and 5.5m±0.5m, respectively. The position between the two signals was also varied proportionally with the length of the central corridor. The initial position of the robot was selected randomly within a 50x45cm rectangular area located at the beginning of the central corridor. The initial orientation of the robot was selected randomly with a uniform distribution in the [-180, 180]° range. Robots were evaluated for 16 or 32 trials, as explained below. Trials lasted up to 1 minute and were stopped as soon as the robot turned in the wrong direction at one of the two junctions.

Complex T-Mazes have already been used in previous research. In particular [32] evolved robots for the ability to navigate in T-Maze environments in multiple trials in which the destination location remained stable. The authors manage to develop robots that were able to "remember" the target location in simple T-Maze but not in double T-Maze environments. [33] evolved robots for the ability to navigate in a triple T-Maze toward 8 alternative destinations. In these experiments, however, the robots were not required to master a delayed response task. Indeed, they received and had access to one of eight different corresponding patterns for the entire duration of each trial.

We used the MarXbot [34] agent, which is a circular robot with a diameter of 17cm equipped with 24 infrared sensors, a rotating scanner sensor and an omnidirectional camera. The experiments were run in simulation by using the FARSA open-software tool that includes an accurate simulator of the robot and of the environment [35]. FARSA has been used to successfully transfer results obtained in simulation to hardware in similar experimental settings (e.g., [17, 36]).

The robots' controllers

Evolving robots are provided with neural network controllers. The sensory layer includes eight sensory neurons that encode the average activation state of eight groups of three adjacent infrared sensors, 6 neurons that encode the average activation of the rotating scanner over 60 degrees, and 8 neurons that encode the percentage of green and blue pixels detected in four 90° sectors of the visual field of the camera. The state of the sensory and motor neurons is normalized in the range [0.0, 1.0] and noise is simulated by the addition of random values selected with a uniform distribution in the range [-0.05, 0.05]. The motor layer includes two motor neurons that encode the desired speed of the two corresponding motors (normalized in the range [-10.0,10.0]cm/s) that actuate the differential driving system of the robot.

The experiments were replicated in three different experimental conditions that varied with respect to the architecture of the robots' neural controller as described below:

(S) Simple: The robots were provided with a simple reactive neural network (that always responded in the same way to the same sensory state) in which the sensory neurons were directly connected to the motor neurons.

(C) Continuous: As in the case of the previous condition the neural network controller included direct sensory-motor connections. In addition, the network included an internal layer with 6 neurons that received connections from the sensory neurons and projected connections to the motor neurons. The internal neurons were continuous, i.e., their output state depended on both the activation received from the incoming connections and on their previous activation state (see [37–38]).

(CR) Continuous Recurrent: The neural network was identical to the previous condition, but the internal neurons were also interconnected through recurrent connections.

The state of the sensors, the neurons and the desired speed of the motors were updated every 50ms. The architecture of the neural network controller was kept fixed.

These experimental conditions were chosen to enable us to verify whether and to what extent the problem had reactive solutions, whether the possibility of integrating sensory-motor information over time into internal neuron states would enable the robots to develop better solutions and whether reactive strategies could coexist with cognitive strategies.

The evolutionary process

The initial population consisted of 20 randomly generated genotypes that encoded the connection weights, biases, and time constants of the neural network controllers of 20 corresponding individual robots. Each parameter was encoded with 8 bits and normalized in the range [–5.0, +5.0] in the case of connection weights and biases and [0.0, 1.0] in the case of time constants of continuous neurons.

Each individual was allowed to generate one offspring, i.e., a copy of the parent genotype in which each bit was mutated with a given probability. Each offspring was evaluated and was used to replace the genotype of the worst parent or was discarded depending on whether or not it outperformed the worst parent.

Each experiment was repeated 10 times by starting with different randomly initialized genotypes. The evolutionary process continued in two consecutive phases of 2,000 generations. During the first 1,000 generations, the mutation rate was set at 2% and the evolving robots were evaluated in 16 trials. From generation 1,001 on the mutation rate was set at 1%. From generation 1501 to 2000 on the individuals were evaluated in 32 trials. In some of the experiments the robots were subjected to position and orientation noise during the second phase, as described below.

The fitness of the individuals was computed by averaging the fitness obtained during the different trials. The fitness of each trial corresponded to the inverse of the distance, within the maze, between the robot and the target destination at the end of the trial. In other words, the robots were rewarded for the ability to approach the appropriate destinations.

The experiments described in this paper can be replicated by downloading and installing FARSA and the associated experimental plugin from "https://sourceforge.net/projects/farsa/" and from "http://laral.istc.cnr.it/cognoffpone/dtmaze.tgz".

Results

In this section we report the results obtained in the different experimental conditions described in Section 2.2 and in additional control experiments described below that were carried out to clarify the role of cognitive offloading in the development of cognitive skills.

Overall the performance of the robots (i.e. the percentage of trials in which the evolved robots reached the correct target destination) did not differ significantly between the three experimental conditions in which they were provided with different neural controllers (see Fig 3). By analyzing the performance of the best robot obtained in each experimental condition (Fig 3) we can see how the best (C) and (CR) robots managed to achieve close to optimal performance in 92.8% and 96.3% of the trials, respectively, while the best (S) robot only achieved sub-optimal performance (i.e. 71.7% success). The performance of the best (C) and (CR) robots did not differ significantly (Fisher Exact Test, $p = 0.308$). The performance of the (S) robot, instead, was significantly worse than the performance of the (C) and (CR) robots (Fisher Exact Test, $p<0.001$). The data were obtained by post-evaluating the best robot from the last generation of each replication in 600 trials.

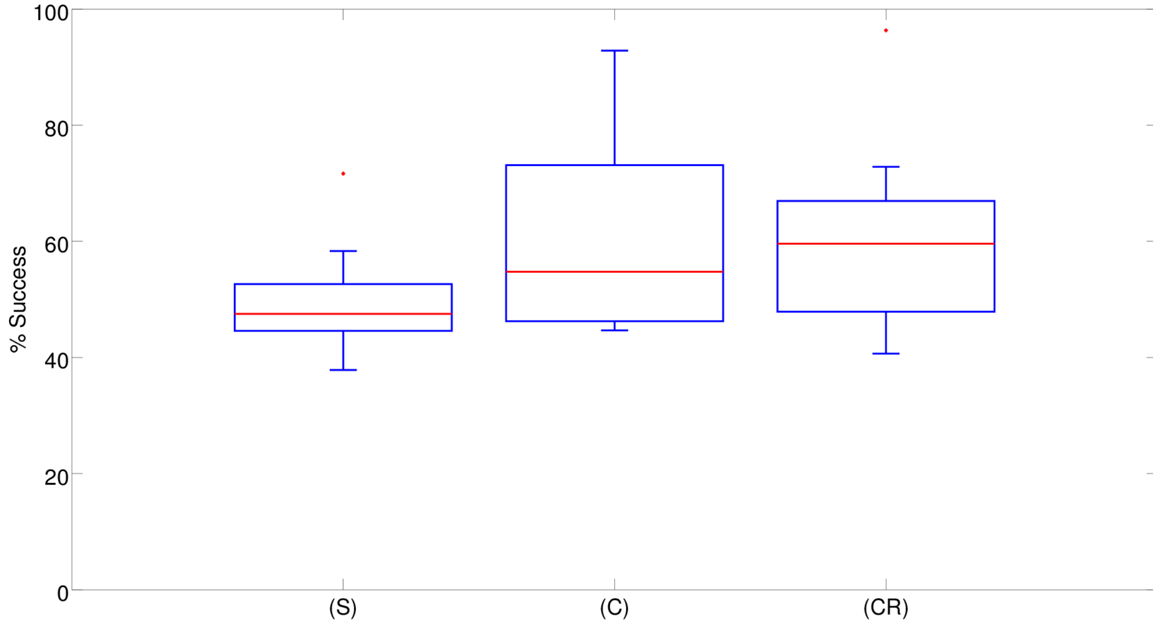


Fig 3. Performance (i.e. percentage of successful trials) of the best 10 robots evolved in the (S), (C) and (CR) experimental conditions in 10 corresponding replications of each experiment. Boxes represent the inter-quartile range of the data and horizontal lines inside the boxes mark the median values. The whiskers extend to the most extreme data points within 1.5 times the inter-quartile range from the box. Circles indicate the outliers.

doi:10.1371/journal.pone.0160679.g003

These results confirm that, at least in our experimental setup, the double T-Maze problem cannot be solved through simple reactive solutions, although the best reactive robot (S) achieved a remarkably high performance (i.e. by navigating toward the correct destination in 71.7% of cases). Moreover, these results indicate that the possibility of integrating sensory-motor information over time through continuous neurons (C) and recurrent connections (CR) into internal states that are used to regulate the way the robots react to sensory stimuli enables the evolving robots to achieve close-to-optimal performances. On the other hand, the fact that close to optimal performances were achieved in a minority of the replications indicates that the task is hard and there is a high probability that evolution remains stuck in sub-optimal regions of the search space.

By inspecting the trajectories produced by the best (S) robot, i.e., by means of a simple reactive controller, we can see how the robot navigated correctly toward all four target destinations most of the time, but erroneously navigated toward the left-left and right-left destinations instead of the left-right destination in several trials (Fig 4, blue trajectories). Surprisingly, this shows that the double T-Maze task can also be solved to a large extent with a reactive solution in which the robot's actions depend only on the current robot's input and in which the robot does not store any internal information regarding previously experienced sensory states. This is achieved by offloading the critical information into the environment (or more precisely into the robot/environmental relationship).

In fact, the behavioral analysis of the best (S) robot indicates that the experienced signals are used to systematically alter the positions assumed by the robot at the end of the central corridor (Fig 4). These positions influence the type of stimuli the robot experiences at the first junction which, in turn, determine whether the robot will turn left or right at the junction. The position of the robot at the end of the first corridor also influences how the turning is realized, i.e., whether the robot produces a tight turn or a wider one, and consequently the position assumed

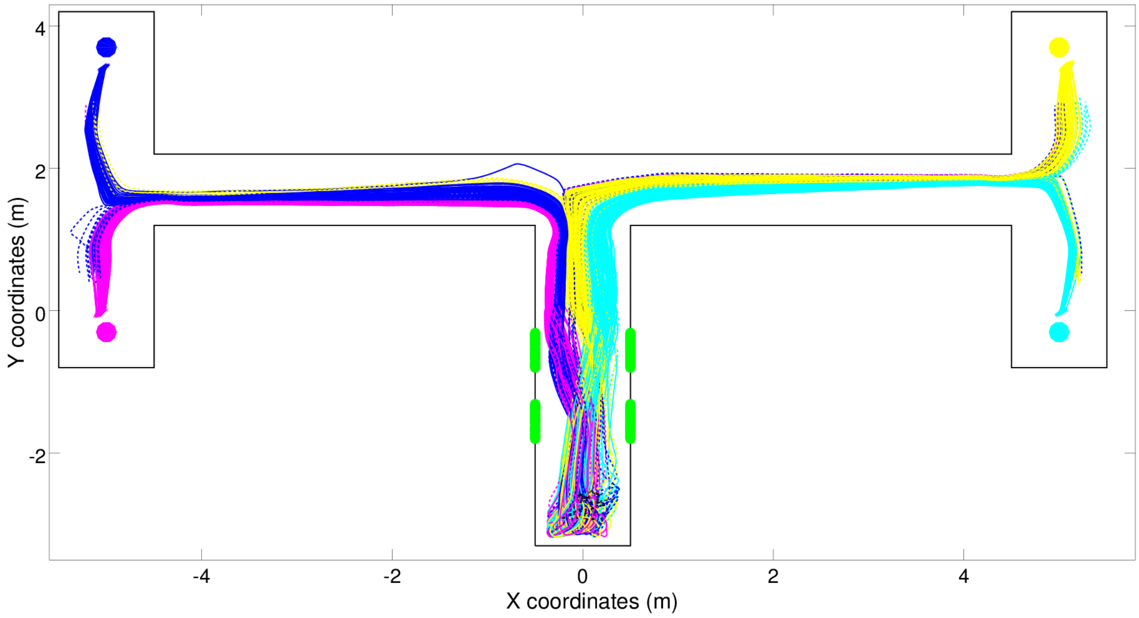


Fig 4. Trajectories produced by the best (S) robot over 300 trials. Full and dashed lines indicate successful and unsuccessful trials, respectively. The color indicates the corresponding target destination (magenta: left-bottom, blue: left-top, cyan: right-bottom, and yellow: right-top).

doi:10.1371/journal.pone.0160679.g004

by the robot in the second corridor. Indeed, after experiencing the right-right signals the robot assumes the right most position at the end of the first corridor and then a right position in the second corridor. By contrast, after experiencing the right-left signals the robot assumes the central position at the end of the first corridor and then a left position in the second corridor. This ability to differentiate the relative position assumed in the second corridor on the basis of the position assumed at the end of the first corridor enables the robots to turn in the appropriate direction also at the second junction. The same things happens when the robots travel towards the other two left destinations.

This interpretation is confirmed by the result of an analysis in which the robot was initially placed at the end of the first corridor with an orientation that systematically varied in the range of [-55, 55]° with respect to the orientation of the first corridor and a position that systematically varied along the x-axis between [-0.4, 0.4]m with respect to the center of the corridor. As shown in Fig 5, in fact, the destination reached by the robot depended primarily on the relative position along the x-axis and secondarily on the orientation of the robot at the end of the first corridor. This means that the robot offloads the information encoding the destination to be reached in its position and orientation. The ineffective behaviors shown in red, which are produced when the robot is initially located near the right wall and oriented toward the right, occur only occasionally in normal conditions (Fig 4) because the robot rarely reaches these positions/orientations when it starts from the beginning of the central corridor.

The fact that the destination reached by the robot depended on the position and the orientation of the robot at end of the first corridor is demonstrated by the fact that the destination the robot will reach at the end of the trial can be predicted with 84% accuracy on the basis of the position and the orientation the robot assumes in the first corridor, 40cm before the first junction. This success rate was obtained by training a feed-forward neural network through a back-propagation algorithm based on a cross-entropy error function [39]. The network included two inputs neurons that encoded the x-position and the orientation of the robot, six hidden

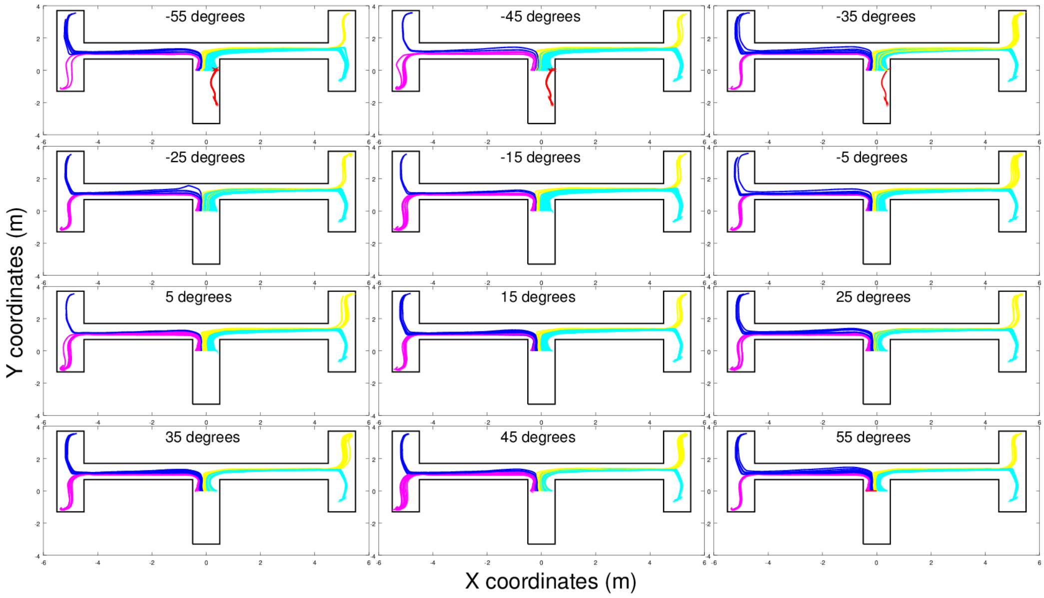


Fig 5. Behavior displayed by the best (S) robot placed initially at the end of the first corridor in different positions (along the x-axis) and orientations. The color of the trajectories corresponds to the destination reached (magenta: left-bottom, blue: left-top, yellow: right-top, and cyan: right-bottom). Trajectories that did not enable the robot to reach any target destination are shown in red.

doi:10.1371/journal.pone.0160679.g005

neurons and four winner-take-all outputs neurons that encoded the four corresponding destinations. The training set consisted of 6,000 position and orientation vectors and 6,000 corresponding destination vectors. The fact that the target destination could not be predicted in all cases can be explained by considering the effects of noise on sensors and motors. Further evidence demonstrating that the target destination reached by the best (S) and (C) robots depended on the position and the orientation assumed by the robots from the end of the first corridor on are reported below.

The four behaviors displayed by this robot (indicated by the trajectories shown in magenta, blue, cyan and yellow in Fig 4) are dynamical processes that arise from the robot/environmental interactions and that converge toward four fixed-point attractors. This can be appreciated by observing the four corresponding basins of attraction in the 2D projection of the phase portrait (Fig 6). These basins of attraction enable the robot to reach four different destinations without varying the way it responds to perceived stimuli (i.e. by using a reactive controller that always responds in the same way to the same stimuli independently from the stimuli experienced before). This can be explained by considering that the way in which the robot reacts to perceptual stimuli and the way in which perceptual stimuli change (as a function of the action performed by the robot and of the characteristics of the local portion of the environment) ensure that the robot keeps moving towards the current destination while remaining in the current basin of attraction. To solve the problem, therefore, the robot only needs to enter into the appropriate basin of attraction in the first corridor while it perceives the green stimuli.

The basins of attraction also ensure that the behavior of the robot is robust with respect to perturbations caused by noise and environmental variations (within limits). This is because typically small alterations in the robot's position and orientation do not cause a switch from one basin of attraction to another and consequently do not alter the robot's destination. Moreover this is also because the effects of small alterations tend to be automatically compensated over time by the convergent nature of the attractors that drive the robot away from the borders

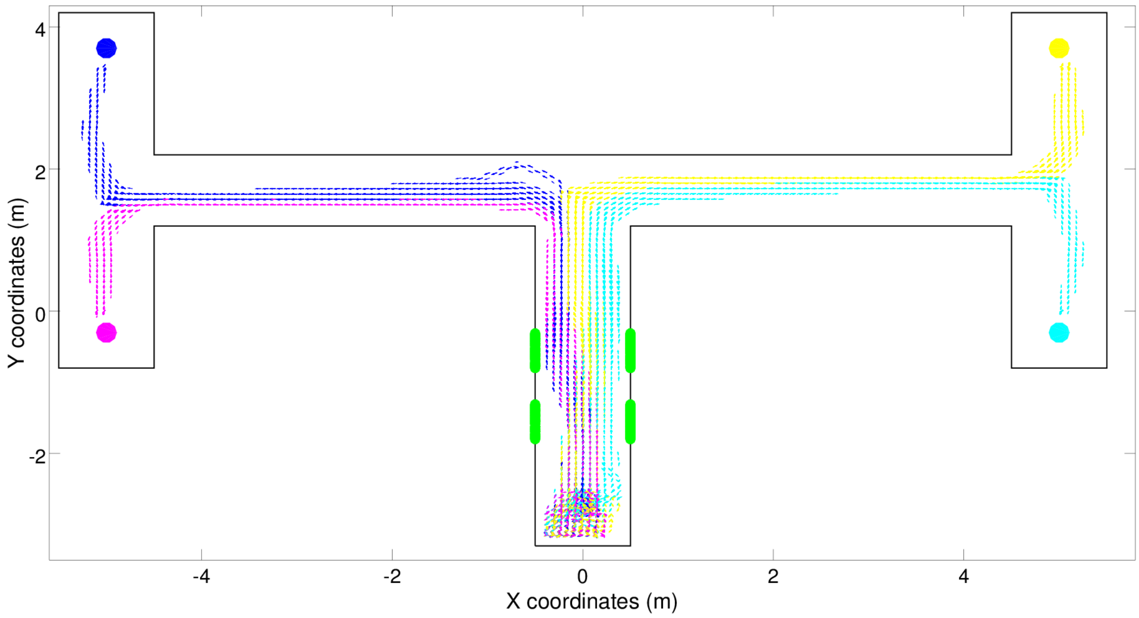


Fig 6. 2D vector field displaying the velocity of the robot in varying positions of the x/y plane for the (S) robot. For the sake of clarity, the vectors are shown only for the positions and orientations reached by the robot in natural circumstances (i.e. the positions and orientation assumed by the robot in 300 trials). The multiple arrows displayed in each 75x75mm position cell indicate how the direction and the magnitude of the velocity vector varies as a function of the different orientations assumed by the robot in the corresponding position.

doi:10.1371/journal.pone.0160679.g006

that separate the different basins of attraction. Note that, as shown in Fig 7, the state space includes three dimensions (x position, y position and orientation). Consequently, the four basins of attraction are separated also in areas in which they seem to overlap from the perspective of the 2D projection shown in Fig 6.

Selection of the appropriate behavior (i.e. the convergence toward the appropriate basin of attraction) is the result of the bifurcation process that occurs in the first corridor and that is

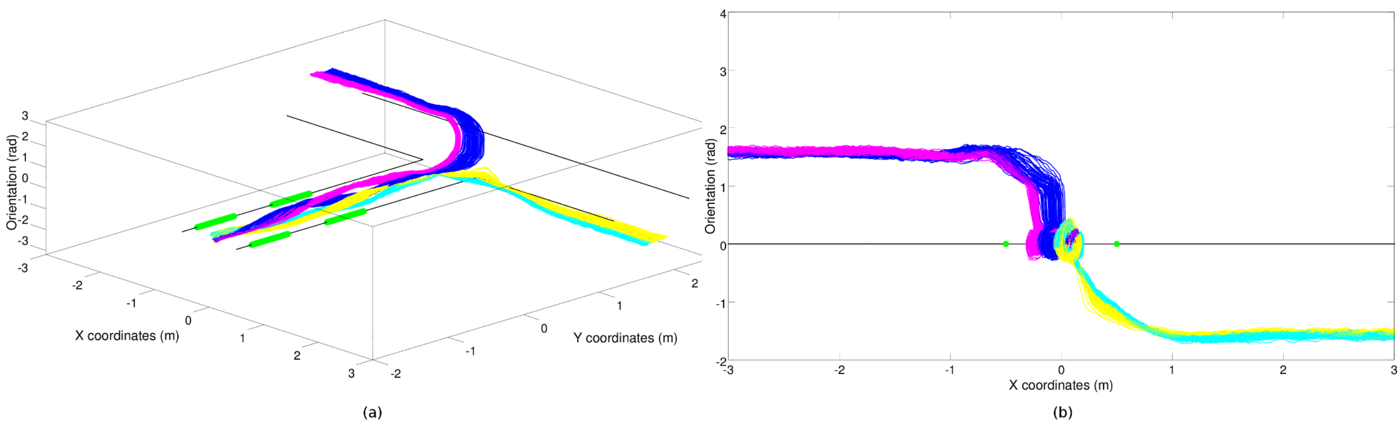


Fig 7. Phase portrait of the robot/environmental dynamics in the case of the best (CR) robot. Data collected during in 300 trials. The vertical axis represents the orientation of the robot in the range [-90, 90]° with respect to the direction of the first corridor. For the sake of clarity the plots refer only to the first-junction portion of the environment. (a) displays a view from which variation on the three dimensions can be appreciated. Instead, (b) shows a 2D orthogonal projection in which one can appreciate only variations along the vertical and horizontal x-axis that correspond to the orientation of the robot and to the position of the robot along the x-axis, respectively.

doi:10.1371/journal.pone.0160679.g007

regulated by perception of the green stimuli. In other words, it is the result of the fact that while the robot travels along the first corridor, it varies its position and orientation on the basis of the perceived green stimuli in a way that ensures that at the end of the first corridor the robot assumes a position and orientation that enable it to enter in the right basin of attraction.

For example, let's consider the trials in which the robots experience the right-right signal (Fig 4, cyan trajectories). In these cases, the robot reacts to the two green stimuli located on the right by moving toward the right side of the corridor. This enables the robot to turn right at the first junction, because it turns right when it encounters a wall ahead and a wall on its right side that is nearer than the wall on its left side, and then to assume a specific position and orientation in the second corridor that enable it to turn right also at the second junction. Thus, the proper movements produced in response to the stimuli perceived in the first corridor ensure that the robot environmental/dynamics will enter into the basin of attraction of the right-right behavior that then guide the robot toward the appropriate destination. During the trials in which the robot experiences the right-left signal, instead, the robot reacts to the signals by moving first right and then left so as to assume a central position within the first corridor (Fig 4, yellow trajectories). This makes the robot turn right at the first junction also in this case. However, the right turn initiated from this position and orientation, drives the robot toward the left side of the second corridor. This in turn enables the robot to then turn left at the second junction. In other words, the position/orientation with which the robot approaches the first junction influences not only whether it turns right or left but also the position/orientation that it takes after the turn in the second corridor, which finally determines whether the robot will turn left or right at the second junction.

There are two reasons for the errors produced by the robot (Fig 4, dashed lines). The first is that in some cases the bifurcation process fails, i.e., the robot is unable to react to the stimuli experienced and, thus, assume the appropriate positions/orientations at the end of the first corridor. Consequently, the robot/environmental system enters into the wrong basin of attraction. This problem particularly affects some of the trials in which the robot experiences the left-right signals. Indeed, in 10.7% of these trials the best (S) robot erroneously navigates toward the right-left destination (Fig 4, dashed lines). This problem occurs when the robot starts from certain specific positions and orientations and/or as a result of noise or when the robot occasionally exits from the right basin of attraction and enters into another, wrong basin of attraction. This typically occurs as a result of noise in areas in which the divergence between two nearby basins of attraction is weak. In the case of the best (S) robot, this second type of problem occurs particularly in the left corridor. Indeed, in this phase the robots traveling toward the left-right destination erroneously enter into the behavioral attractor of the left-left destination in 46.7% of the left-right trials (Fig 4, dashed lines).

The behaviors displayed by the best (C) and (CR) robots are qualitatively similar to those displayed by the best (S) robot, see Fig 8. Indeed, also the robot/environmental dynamic of these robots is characterized by four fixed-point attractors (Fig 9). Moreover, the trajectories of these robots also bifurcate in the first corridor to ensure that the robot/environmental dynamic enters into the appropriate basin of attraction.

However, these robots make far fewer errors during the bifurcation phase than the (S) robot thanks to their ability to converge toward similar positions and orientations while they move along the first part of the first corridor (see Fig 8). Indeed, the variability along the x-axis of the positions assumed by (C) and (CR) robots one meter before the green stimuli is significantly lower than the variability of positions assumed by (S) robots (F-Test $F(299,299) = 3.114$, $p<0.001$ and $F(299,299) = 2.845$, $p<0.001$, respectively). By assuming a relatively un-variant position and orientation before they perceive the green stimuli, (C) and (CR) robots manage to achieve a more reliable bifurcation process than (S) robots.

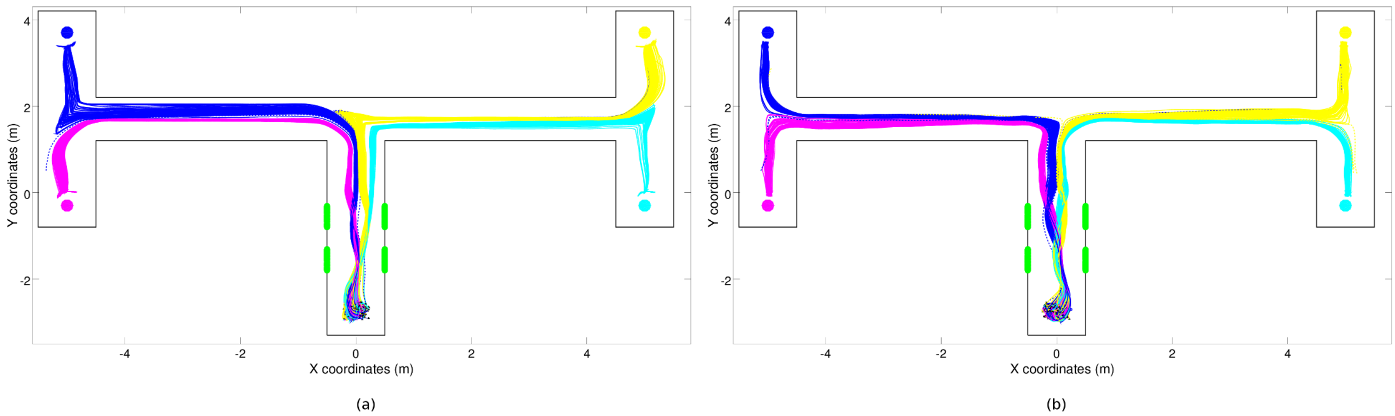


Fig 8. Trajectories produced by the best (C) and (CR) robots in 300 trials, (a) and (b) pictures respectively. Full and dashed lines indicate successful and unsuccessful trials, respectively. The color indicates the corresponding target destination (magenta: left-bottom; blue: left-top, cyan: right-bottom, and yellow: right-top).

doi:10.1371/journal.pone.0160679.g008

Also, the errors that occur when the robots exit from their current basin of attraction and enter into another basin of attraction are reduced in (C) and (CR) robots. This is achieved through the synthesis of attractor basins, which produces a greater separation among the trajectories targeted toward different destinations that are produced by (C) and (CR) robots (Fig 8) than among the trajectories produced by the (S) robot (Fig 4).

Finally, in some cases the best (CR) robot also displayed the ability to re-enter into the basin of attraction in which it was previously located when it erroneously moved to another basin of attraction. This enables the robot to recover from some of the errors of this type caused by

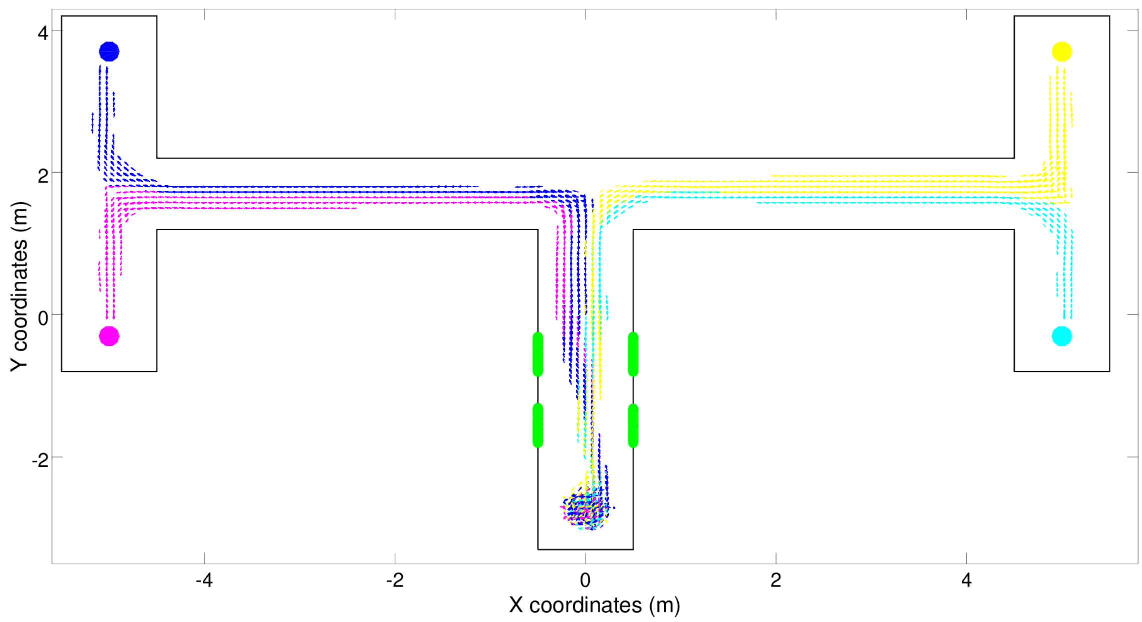


Fig 9. Vector field displaying the velocity of the robot in varying positions of the x/y plane for the (CR) robot. For the sake of clarity, the vectors are shown only for the positions and orientations reached by the robot in natural circumstances (i.e. the positions and orientations assumed by the robot in 300 trials). The multiple arrows displayed in each 75x75mm position cell indicate how the direction and the magnitude of the velocity vector varies as a function of the different orientations assumed by the robot in the corresponding cell and/or as a function of the robot's internal states.

doi:10.1371/journal.pone.0160679.g009

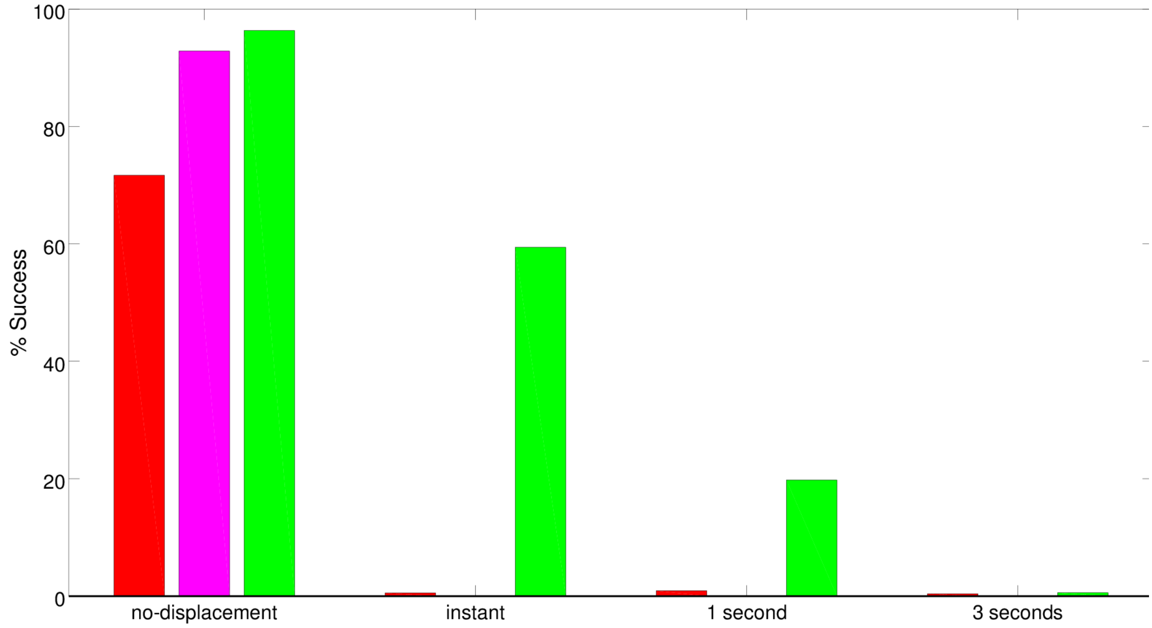


Fig 10. Performance displayed by the best (S), (C), and (CR) robots during a post-evaluation test in which they were randomly displaced from their current basin of attraction into another one. The red, magenta and green bars represent the performances of the best (S), (C) and (CR) robots, respectively. Each bar represents the percentage of trials in which the robots were able to reach the appropriate destination despite the displacement. The "instant" condition refers to a situation in which immediately after the displacement the robots were allowed to move normally. The "1 second" and "3 seconds" conditions refer to a situation in which the robots were unable to move for 1 and 3 seconds, respectively. The "no-displacement" condition refers to a normal situation in which the robots were not subjected to displacements. During each trial the robot was displaced when it reached an imaginary line located 40cm before the first junction, 40cm after the first junction, in the middle of the second corridor or 40cm before the second junction. Data were collected from 2400 trials.

doi:10.1371/journal.pone.0160679.g010

noise. This is demonstrated by the results collected during a post-evaluation test in which the robot was systematically displaced from its current basin of attraction to another one. The basin of attraction in which the robot is displaced is chosen randomly from among the other available ones, i.e., from the other three alternative basins of attraction before the first T-junction or between the only alternative attractor after the first T-Junction. The post-evaluation test was repeated in three conditions in which the robot was allowed to move normally and in which it was blocked for 1 or 3 seconds after displacement. Analysis of the results indicates that the best (CR) robot was able to recover from this type of displacement in 60% of cases in which it was allowed to move normally after the displacement and in 20% of cases in which it was blocked in the displaced position and orientation for 1 second (Fig 10). The best (S) and (C) robots, instead, were able to recover from displacements only in a negligible percentage of cases (Fig 10). None of the robots were able to recover from displacements after being blocked in the displaced position and orientation for 3 seconds.

Overall the data reported in this section indicate that all robots offloaded information concerning the stimuli they experienced in the position and orientation they assumed from the end of the first corridor on and used this information to move toward the appropriate destination and to preserve the relevant information (i.e. to maintain a specific relative position and orientation with respect to the environment). (C) and (CR) robots also exploited the possibility of integrating sensory-motor information over time to regulate their motor behavior in the very first portion of the first corridor so as to reduce the variability with which they reached the green stimuli. Moreover, (C) and (CR) robots also exploited the possibility of integrating

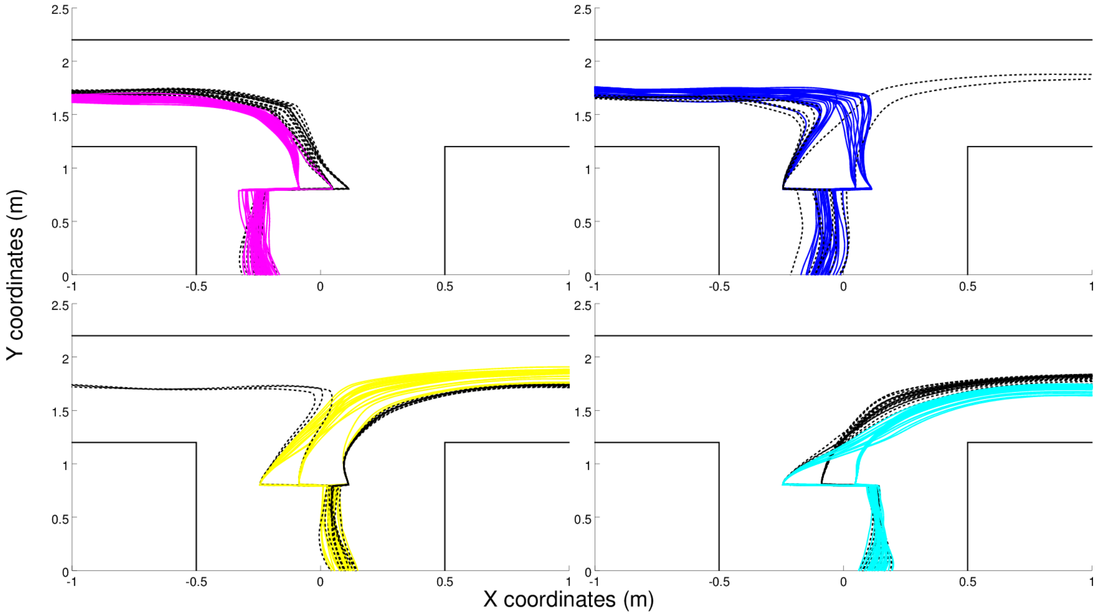


Fig 11. Trajectories displayed by the best (CR) robot at the first T-junction during a post-evaluation test in which the robots were displaced into one position and orientation located within one of the other three basins of attraction. The displacement was performed when the robot reached a distance of 40cm from the first junction. After the displacement the robot was allowed to move immediately (i.e. it was not blocked). Data were collected from over 200 trials, 50 for each combination of green stimuli. The colors indicate the target destinations (magenta: left-bottom, blue: left-top, cyan: right-bottom, and yellow: right-top). The black trajectories indicate the trials in which the robot reached a wrong destination, i.e. was unable to re-enter the correct basin of attraction after the displacement.

doi:10.1371/journal.pone.0160679.g011

sensory-motor information over time to partially filter out the effect of noise affecting their sensors and motors and to better separate the trajectories produced while they navigated toward different target destinations.

The destination reached by (S) and (C) robots was determined by the position and the orientation assumed by the robots from the end of the first corridor on and was not affected by the type of stimuli experienced previously. Indeed when these robots were displaced from their current position inside a certain basin of attraction into a position and an orientation located in a different basin of attraction, they navigated towards the target destination corresponding to the second basin of attraction in 98% of cases. The destination reached by (CR) robots, instead, also depended on the state of the internal neurons that encode information about previously experienced sensory states. Indeed, displaced (CR) robots navigated toward the destination corresponding to the basin of attraction in which they were located before the displacement in 60% of cases. They managed to compensate the effect of the displacement by re-entering the basin of attractions in which they had been previously located. They navigated toward wrong destinations only in the remaining 40% of cases (see Fig 11).

Limiting cognitive offloading does not promote but rather prevents the synthesis of effective solutions

Here we report a series of experiments carried out to verify whether the discovery of sub-optimal reactive strategies that rely on cognitive offloading prevents the discovery of better cognitive solutions. In other words, we verify the hypothesis that cognitive offloading constitutes a sort of shortcut that enables evolving individuals to improve their adaptive ability up to a certain level through the utilization of solutions that are parsimonious from the perspective of the

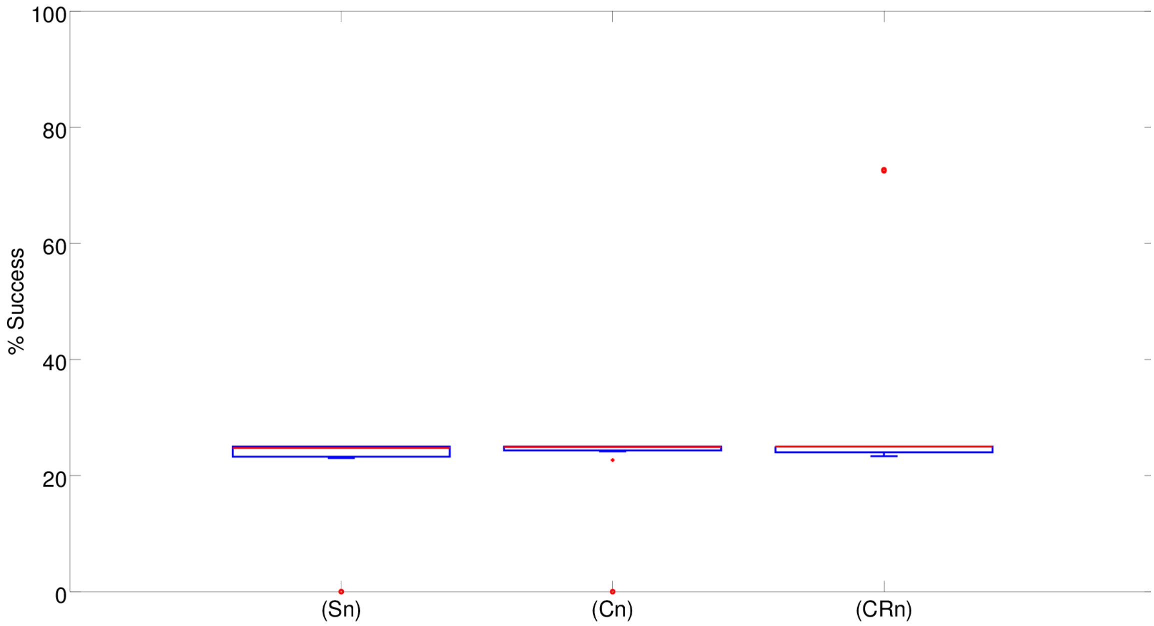


Fig 12. Performances obtained in the control experiment with narrow corridors. The boxplots show the percentage of successful trials carried out by the best 10 robots evolved in the (S), (C), and (CR) experimental conditions in 10 corresponding replications of the experiment. Boxes represent the inter-quartile range of the data. 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.

doi:10.1371/journal.pone.0160679.g012

control system of the robot but that prevent the discovery of more complex and effective strategies. To achieve this objective we analyzed the solutions found by evolving robots in situations in which the possibility of relying on cognitive offloading was reduced or prevented.

One way to reduce the possibility of relying on cognitive offloading in the case of our experiments was to drastically reduce the width of the corridors. As we have seen, in fact, robots offload information concerning the type of green stimuli they have experienced by assuming different relative positions/orientations inside corridors. The utilization of narrow corridors severely restricts the possibility of carrying out this type of offloading.

Analysis of the results obtained in a series of control experiments in which the width of the corridors was set to 29cm only indicates that, as expected, the use of highly constrained environmental conditions prevents the exploitation of cognitive offloading, i.e., the evolution of solutions analogous to that described in the previous section (results not shown). However, analysis of the performance of the robots evolved in this condition indicates that the elimination of solutions based on cognitive offloading does not lead to effective solutions. In fact, it causes a drastic reduction of the robots' performance with respect to the normal condition (Fig 12). This implies that the elimination of solutions relying on cognitive offloading does not facilitate the evolution of alternative cognitive solutions. In other words, the lack of evolution of effective cognitive solutions cannot be explained simply by the availability of cognitive offloading "shortcuts".

Indeed, in most of the replications the best (Sn), (Cn) and (CRn) evolved robots were able to navigate correctly to only one of the four destinations and consequently succeeded in about one fourth of the trials. The evolved robots managed to navigate to one of the four correct destinations in most of the trials. Consequently, the impact of the errors that occurred when the robots remained stuck, navigated erroneously back toward the central corridor or crashed into obstacles was marginal. We used (Sn), (Cn) and (CRn) to indicate the robots evolved in the narrow corridor condition. In only two replications, the best (CRn) robots managed to navigate correctly toward three out of the four destinations by succeeding in about three fourths of the trials (Fig 12). The performance of these two robots was similar to that achieved by the best (S) robot (Fisher exact test, $p = 0.747$ and $p = 0.797$), which could only rely on reactive strategies (Fig 3). However it was significantly worse than that of the better (C) and (CR) robots, (Fisher exact test, $p<0.001$). Overall, the performance of the robots that evolved in these control experiments (Fig 12) was significantly worse than the performance obtained in the standard experiments (Fig 3) (Mann-Whitney U, $p<0.001$ for (S) and (C), and $p = 0.01$ for (CR)).

Another mechanism that can be used to discourage evolving robots from relying on cognitive offloading is to randomly vary the position and orientation of the robots while they travel in the maze. To investigate the effect of this type of perturbation we carried out a series of control experiments in a standard maze in which every 50ms the position and the orientation of the robot was perturbed with a 15% probability. The perturbations were created by displacing the robot to the left or the right of a distance d and by varying the orientation of the robot by an angle a, where d and a are selected randomly with a uniform distribution within the range [3, 9]cm and [-15, 15]°, respectively. This type of perturbation drastically reduces the utility (usefulness) of offloading information in the relative positions and orientation of the robot in the environment. Once again, the hypothesis under verification is whether the introduction of a constraint that discourages the development of cognitive offloading solutions will favor the development of alternative, and possibly better, solutions. We used (Sp), (Cp) and (CRp) to indicate the robots evolved in the standard maze subjected to position and orientation perturbations during evolution.

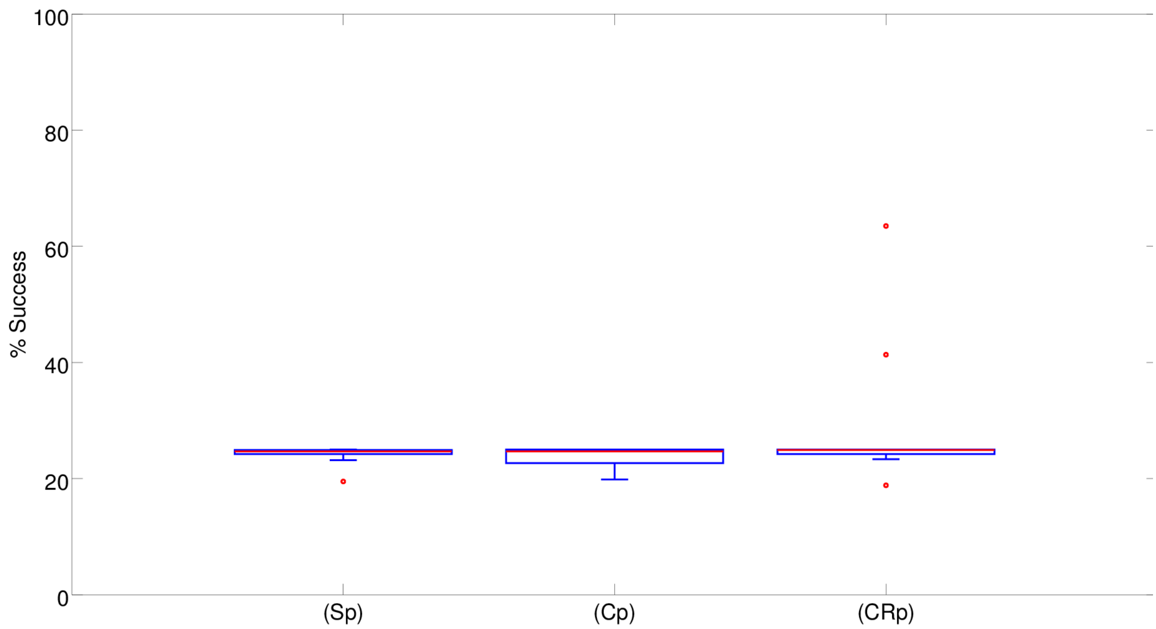


Fig 13. Performance obtained by robots that were subjected to position and orientation perturbations in all trials during 4000 generations. Data were obtained by post-evaluating the robots in a normal condition in which they were not subjected to perturbations. The boxplots show the percentage of successful trials carried out by the best 10 robots evolved in the (Sp), (Cp), and (CRp) experimental conditions in 10 corresponding replications of the experiment. Boxes represent the inter-quartile range of the data. 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.

doi:10.1371/journal.pone.0160679.g013

The fact that this form of perturbation drastically reduces the usefulness of cognitive offloading is demonstrated by the fact that the performance of (Sp) robots, which can only rely on reactive strategies, drops to very low levels (Fig 13). The introduction of perturbations, however, causes a significant drop in performance with respect to the experiments without perturbations also in the case of (Cp) and (CRp) robots (Mann-Whitney U, $p<0.001$) (Fig 13). Notice that Fig 13 displays the results of a post-evaluation test in which the robots are evaluated in the absence of position and orientation perturbations.

Also in this case, therefore, the introduction of constraints that discourage the development of reactive solutions relying on cognitive offloading does not promote the evolution of effective cognitive solutions but rather prevents the possibility of synthesizing good solutions of any kind.

As we will show in the next section, this negative result is not due to the impossibility of generating solutions that are able to compensate the effect of position and orientation perturbations. Indeed, evolving robots can solve the navigation problem in a close to optimal manner and can neutralize to a good extent the effect of position and orientation perturbations, providing that the constraints which discourage the development of cognitive offloading strategies are not too strong.

The acquisition of reactive strategies promotes the evolution of cognitive capabilities

As demonstrated in the previous section, preventing or severely limiting the possibility of developing strategies based on cognitive offloading prevents the development of effective solutions. In this section, we demonstrate how the acquisition of reactive strategies promotes the evolution of cognitive solutions that enable the robots to accomplish their task also in conditions that cannot be mastered appropriately only by using cognitive offloading strategies.

To verify this hypothesis we carried out a new series of experiments in which we weakened the constraints that discourage the utilization of cognitive offloading and increase the demand for cognitive solutions. This was carried out by subjecting the robots to position and orientation perturbations in only half of the trials (Fig 14, CRp2 condition) or by subjecting the robots to perturbations only during the second phase of the evolutionary process, i.e., from generation 2001 to 4000 (Fig 14, CRp3 condition). Notice that the Fig 14 displays the results of a post-evaluation test in which the evolved robots were not exposed to position and orientation perturbations.

The performances of the robots that were subjected to perturbations in only half of the trials are significantly better than the performances of the robots that experienced perturbations in all trials (Fig 14, CRp2 and CRp, Mann-Whitney U, $p = 0.004$). The performances of the robots that were subjected to perturbations during the second phase of the evolutionary process, instead, did not differ significantly from those of the robots that were subjected to perturbations during all generations (Fig 14, CRp3 and CRp, Mann-Whitney U, $p = 0.36$).

The fact that the robots evolved in the CRp2 condition relied on effective cognitive mechanisms can be demonstrated by post-evaluating the best robot in a control condition in which it was systematically displaced from its current position and orientation into another position and orientation located in a different basin of attraction through the same procedure described in Section 3. As can be seen, unlike the best CR robot, the best CRp2 robot is able to recover from the displacements in most cases, also in the condition in which it is blocked for three seconds after being displaced (Fig 15). The robot compensates for the effect of displacement by re-entering into the basin of attraction in which it was located before the displacement. This is carried out by storing in its internal states information encoding the type of basin of attraction

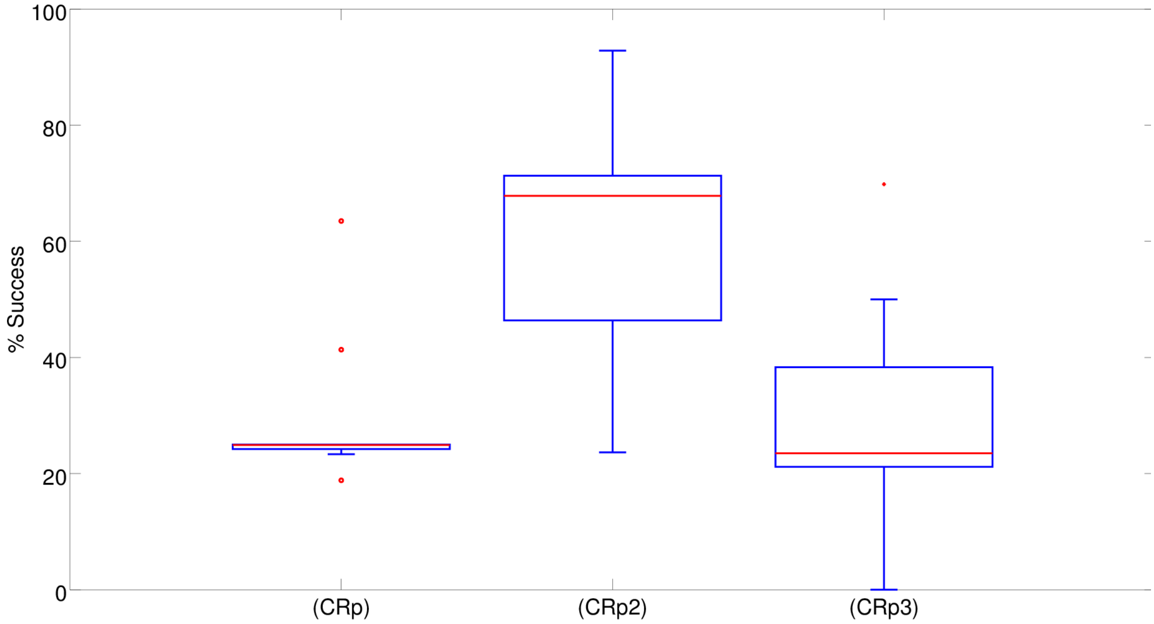


Fig 14. Performances obtained in experiments in which the robots were subjected to position and orientation perturbations in all cases (CRp), during half of the trials (CRp2) and during the second evolutionary phase (CRp3). Data obtained by post-evaluating the robots without position and orientation perturbations. The boxplots show the percentage of successful trials carried out by the best 10 robots evolved in each of the experimental conditions in 10 corresponding replications of the experiment. Boxes represent the inter-quartile range of the data. 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.

doi:10.1371/journal.pone.0160679.g014

in which it is located and by using this information to re-enter into the previous basin of attraction, as shown in Fig 11. The best CR and CRp2 robots make use of the internal states to neutralize the effects of this type of displacement (see Figs 10 and 15). However, the best CRp2 displays much better ability. This is not surprising, because CRp2 robots were subjected to position and orientation perturbations during evolution. In normal conditions, i.e., without displacements or perturbations, the performances of the best CR and CRp2 controllers do not differ statistically (Fisher Exact Test, $p = 0.308$).

The fact that the development of cognitive offloading strategies supports the development of effective cognitive abilities, such as those displayed by the best CRp2 robot, is also demonstrated by the analysis of the course of the evolutionary process in the case of the best replication of the experiment. In fact, the post-evaluation of the best robots in every 50 generations indicates that the development of the ability to master the trials not affected by displacements precedes the development of the ability to also master the trials subjected to displacement (Fig 16). The performances in the two conditions differed significantly from generation 1500 to 1800 with the exception of generation 1700 (Fisher Exact Test, $p<0.05$). We focused on this 500 generations period because in the case of the best replication this is the phase in which most progress occurs. Thus, the development of a strategy that operates primarily on the basis of cognitive offloading and that enables the robots to handle the navigation task in the normal condition but not in the condition with displacements supports the development of a hybrid strategy that relies also on internal states and that enables the robots to also master the trials affected by displacements.

Incidentally, the internal neurons of the (CR) and (CRp2) robots are somewhat similar to the hippocampal cells of rodents located in maze environments that encode information about the location of the animal in the maze and about the trajectory the animal is performing and will perform to reach the target destination [40–42]. However, a detailed analysis of the relation with these neurophysiological findings is outside the scope of this paper.

At this point we will try to explain why evolving robots: (i) always develop cognitive offloading strategies, (ii) are unable to develop effective strategies relying exclusively on internal states even in conditions in which the possibility of using cognitive offloading is completely ruled out, (iii) are able to extract internal states encoding target destinations that are maintained over time and are used to compensate the effects of position and orientation perturbations.

We believe that to explain these results we have to appreciate the full complexity of the task. In particular, we must consider that the problem involves both the ability to travel in the maze environment by avoiding collisions and inversions in the direction of motion and the ability to turn in the appropriate direction at junction areas. The most straightforward way of navigating in the maze and avoiding collisions is to use a reactive strategy that regulates the direction of the robot on the basis of the current state of the infrared sensors. In other words, the ability to navigate in the maze environment necessarily requires use of a reactive strategy. The capacity to turn in the appropriate direction at junction areas, instead, can be obtained either by exploiting primarily external cues generated through cognitive offloading or by exploiting primarily internal cues generated by integrating sensory-motor information over time. Using a mixed strategy that operates both on the basis of reactive rules for the purpose of navigation and on the basis of more complex rules for the purpose of decision-making at junctions necessarily requires the incorporation of mechanisms that sort out the conflicts arising between reactive and non-reactive control rules.

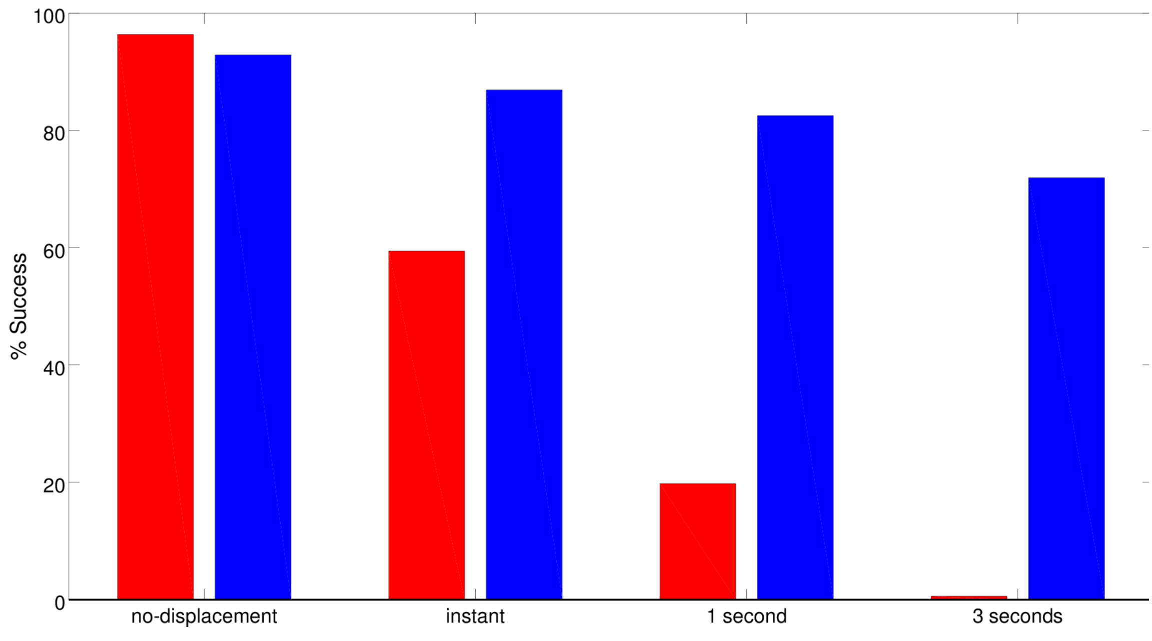


Fig 15. Performance displayed by the best (CR) and (CRp2) robots during a post-evaluation test in which the robots were systematically displaced into a position and orientation located in a different basin of attraction (red and blue bars, respectively). Each bar represents the percentage of trials in which the robots were able to reach the appropriate destination despite the displacement. The "instant" condition refers to a situation in which the robots were allowed to move normally immediately after the displacement. The "1 second" and "3 seconds" conditions refer to situations in which the robots were blocked for 1 and 3 seconds, respectively, after the displacement. The "no-displacement" condition refers to a normal situation in which the robots were not subjected to displacements. The robots were displaced once during each trial. The displacement occurred when they reached an imaginary line located 40cm before the first junction, or 40cm after the first junction, or in the middle of the horizontal corridor, or 40cm before the second junction. Data collected in 2400 trials.

doi:10.1371/journal.pone.0160679.g015

Fig 16. Performance displayed by the best CRp2 during the course of the evolutionary process in a no-displaced and displaced condition (black and red lines, respectively). In the latter condition the robot was allowed to move immediately after the displacement. Data were collected every 50 generations from generation 1500 to 2000 and averaged over 600 trials in the case of the no-displacement condition and 2400 trials in the case of the displaced condition. The displacement occurs when the robot reaches an imaginary line located 40cm before the first junction, or 40cm after the first junction, or in the middle of the horizontal corridor, or 40cm before the second junction.

doi:10.1371/journal.pone.0160679.g016

As an example of conflict, let's consider the case of a robot that decides which direction to take at a junctions on the basis of its internal state. Moreover, let's imagine that the robot is located at the beginning of the first junction and that is traveling toward one of the two left destinations. During the first and the third part of the junction negotiation the robot should not turn too sharply left to avoid colliding with the left walls. During the central part of the junction negotiation, instead, the robot should definitely turn left following the indication coming from its internal state. The conflict arising between obstacle avoidance and the decision-making rules that control the direction of turns at junctions could be solved by regulating the relative importance of the two rules on the basis of the relative position of the robot within the junction. However, the robot does not have access to position information. It only perceives the proximity of nearby obstacles. This implies that the robot would be forced to determine the relative importance of the two alternative control rules on the basis of the indirect, noisy, and incomplete information provided by its sensors. Consequently, this implies that the probability that the robot will fail as a result of ineffective regulation is not negligible.

Instead, the behavioral attractor strategies displayed by evolved robots do not require differentiating the way in which the robots react to sensory states experienced in different portions of the environment. Indeed, once the robots enter into the right basin of attraction, they just need to keep moving on the basis of simple reactive rules in both corridors and junctions. One reason why evolving robots always select cognitive offloading solutions is that these strategies are more robust, less prone to errors with respect to strategies in which turning decisions are made on the basis of internal states.

The second reason why evolution always converges toward cognitive offloading strategies is that preparatory actions that anticipate in part the execution of the required behavior are adaptive and therefore tend to be selected. This implies that the individuals that anticipate the movement toward the left or the right side of the corridor during the trials in which they should turn toward that side at the first and/or at the second junctions tend to be selected. This produces a progressive anticipation of the time when the turning actions are initiated that ultimately leads to a situation in which they are initiated up to the point when the robot perceives the green stimuli. This in turn eliminates the need to extract and use internal states that encode information about previously experienced stimuli. Indeed, we might say that in the evolved robots the left or right turning behaviors produced at the first and second junctions are initiated already during the first half of the first corridor when the robot perceives the position of the green stimuli. In other words, by anticipating action execution through preparatory actions the robots manage to transform a time delay task into a simpler problem that does not include any temporal offset between the perception of stimuli and the initiation of the action afforded by the stimuli. Overall this implies that the selection of cognitive offloading strategies of this type is inevitable, at least in the case of the double T-Maze experimental setting.

The tendency to anticipate behaviors through the execution of preparatory actions provides two advantages: it enables the execution of smoother transitions between behaviors (i.e. between the navigation behavior performed within the corridors and the turning behavior performed within the junctions), and it enables reducing and/or eliminating the time delay between the moment when the stimuli affording a given behavior are experienced and the moment when the behavior afforded by the stimuli is executed.

As we have showed above, however, the ability to solve the time-delay problem through preparatory actions does not necessarily prevent development of the ability to extract information encoding the type of basin of attraction in which the robot is currently located or the green stimuli that the robot experienced and the development of an ability to use these internal states to navigate to the appropriate destination. Indeed, as we have seen, (CR) robots display the ability to re-enter into the basin of attraction in which they were previously located after being displaced into another wrong basin of attraction. In the case of robots that are subjected to a moderate level of position and orientation perturbation (CRp2), this cognitive capability is so effective that it enables the robots to compensate for the effect of the displacement in most cases, even when the robots are blocked for three seconds after the displacement. This type of redundant solution enable these robots to exploit the advantage of preparatory actions and to avoid the problems caused by noise and by position perturbations.

Anticipation is a widespread phenomena in sequential motor control. The preparatory actions that support the realization of effective grasping behaviors are an example of this. These preparatory actions involve appropriate modification of the posture of the hand performed during execution of the reaching behavior that precedes the grasping action [43]. A second example is constituted by the co-articulatory movements produced by sign language interpreters engaged in fingerspelling. In fact, the posture of the hand that is used to indicate a letter is influenced by the posture that the hand should assume later to indicate the following letters [44].

Discussion

In this paper we investigated whether the development of reactive solutions promotes or prevents the development of cognitive solutions to problems in which reactive control policies enable the achievement of sub-optimal performance only. For this purpose we carried out a series of experiments in which evolving robots had to solve a time-delay task in a double T-maze environment in which the destination to be reached depended on the stimuli perceived by the robot during the initial phase of the navigation. The problem chosen is qualitatively similar but more complex than the tasks used in previous studies that investigated the evolution of cognitive capabilities [19][27–31]. The additional complexity is that our task involves a greater number of different destinations, requires making two subsequent stimuli-dependent decisions, involves a longer time delay between the moment in which the robot experiences the stimuli and the moment in which it should make the corresponding turning decisions, and involves variations affecting the initial position and orientation of the robot, the position of the stimuli and the size of the environment.

Analysis of the experiment in which the robots were provided with reactive controllers confirmed that the problem does not allow for optimal, or close-to-optimal, reactive solutions (Fig 3). Surprisingly, however, reactive robots managed to solve the task to a large extent. Analysis of the solutions discovered by the evolving robots indicate that this is achieved by exploiting cognitive offloading. Indeed, the evolved robots display an ability to extract critical states, store these states in the robot/environmental relations and regulate their behavior on the basis of their relative position in the environment (Fig 4).

Analysis of the robots provided with richer neural controllers indicated that the possibility of storing internal states enables the evolving robots to achieve close-to-optimal performance, i.e., to achieve better performance with respect to reactive robots (Fig 3). Analysis of these experiments indicates that cognitive offloading also plays a key role in these robots (Figs 8 and 10). The achievement of better results is due to the development of additional cognitive capabilities that overcome the limitations displayed by robots that operate on the basis of reactive controllers only. The results obtained thus indicate that the development of the reactive strategies based on cognitive offloading does not prevent the development of solutions that rely on cognitive capabilities.

This conclusion is further supported by results obtained in other experiments in which the usefulness of cognitive offloading was reduced or eliminated by using an environment formed by narrow corridors or by subjecting evolving robots to position and orientation perturbations. As expected, the robots evolved in these conditions relied less or not at all on cognitive offloading. This, however, did not enable the robots to discover alternative strategies for solving the task. This simply led to the robots displaying rather ineffective solutions (Figs 12 and 13). Overall, this indicates that the elimination of cognitive offloading does not promote but rather prevents the synthesis of effective solutions.

Finally, we demonstrated how the acquisition of reactive strategies promotes the evolution of cognitive strategies, or better of hybrid strategies that include both cognitive offloading and cognitive mechanisms. These hybrid strategies enable the robots to navigate toward the appropriate destination also after being displaced into another basin of attraction and also after being blocked there for three seconds (Figs 14 and 15). This type of solution was obtained by weakening the constraints that reduce the usefulness of cognitive offloading (i.e. by perturbing the position and orientation of the robot in only half of the trials). This, in fact, creates the appropriate demand for the development of cognitive abilities without preventing the development of cognitive offloading strategies.

Overall these results indicate that reactive strategies relying on cognitive offloading do not necessarily constitute a dead end that might retard or prevent the evolution of better cognitive strategies. On the contrary, they constitute an important component of effective solutions and can co-exist and support the development of complementary cognitive capabilities. For results collected in human subjects that indicate how cognitive offloading can favor the acquisition of abstract concepts see [45].

The importance of the incremental nature of the evolutionary process and of the acquisition of reactive strategies in the synthesis of better cognitive strategies is further demonstrated by the analysis of the course of the evolutionary process. Indeed, the evolution of the cognitive offloading ability precedes the evolution of the cognitive capabilities that use internal states to determine the travel destination (Fig 16).

One question that remains open is whether solutions that are purely cognitive, i.e., that do not also rely on cognitive offloading, exist and can be discovered. We believe that the existence of this type of solution cannot be taken for granted, at least in the context of the domain considered in this paper.

In general, analysis of the characteristics of evolved strategies suggests that reactive and cognitive components of control policies should not be considered as neatly separable entities performing well-differentiated functions. Regulation of the robot's actions performed on the basis of internal states should always be integrated with regulation of the robot's actions, which is carried out on the basis of currently perceived states. In the context of our experiments this implies that the efficacy of a cognitive component of the robot policy that determines the turning direction at T-junctions on the basis of internal states strongly depends on the way the robot reacts to perceived stimuli independently of the value of its internal states, and vice versa. Moreover, use of a cognitive offloading strategy for the purpose of storing information about previously experienced stimuli does not necessarily conflict with use of internal states that have the same function. On the contrary, the combined use of two alternative mechanisms with different characteristics for achieving the same function might provide advantages. If we embrace a less simplified view of the relation between reactive and cognitive components we see fewer reasons to expect interferences and more reasons to expect synergies, like those we found in our experiments.

Acknowledgments

This work is partially funded by CAPES through the Brazilian program science without borders. We thank Giovanni Pezzulo and the anonymous reviewers for useful comments.

Author Contributions

Conceptualization: JTC SN.

Data curation: JTC SN.

Formal analysis: JTC SN.

Funding acquisition: JTC SN.

Investigation: JTC SN.

Methodology: JTC SN.

Project administration: JTC SN.

Resources: JTC SN.

Software: JTC SN.

Supervision: JTC SN.

Validation: JTC SN.

Visualization: JTC SN.

Writing - original draft: JTC SN.

Writing - review & editing: JTC SN.

References

1. Varela F., Thompson E.T., & Rosh E. (1991). The Embodied Mind. Cambridge, MA: MIT Press.

2. Thelen E., & Smith L. (1994). A Dynamic Systems Approach to the Development of Cognition and Action. Cambridge, MA: MIT Press.
3. Port R.F., & van Gelder T. (1995) Mind as Motion. Cambridge, MA: MIT Press.
4. Pfeifer R., & Bongard J. (2007). How the Body Shapes the Way We Think: A New View of Intelligence. Cambridge, MA: MIT Press.
5. Clark A. (1999). An embodied cognitive science? Trends in Cognitive Sciences (3:) 9: 345–351.
6. Steward J., Gapenne O., and Di Paolo E.A. (2010). Enaction: Toward a New Paradigm for Cognitive Science. Cambridge, MA: MIT Press.
7. Clark A. (1997). Being There: Putting Brain, Body and World Together Again. Cambridge MA: MIT Press.
8. Keijzer F. (2002). Representation in dynamical and embodied cognition. Cognitive Systems Research, 3: 275–288.
9. McBeath M, Shaffer D., & Kaiser M. (1995). How baseball outfielders determine where to run to catch fly balls. Science 268: 569–573. PMID: 7725104
10. Nolfi S., Bongard J., Floreano D., & Husband P. (in press). Evolutionary Robotics, in Siciliano B., Khatib Oussama (eds.), Handbook of Robotics, II Edition. Berlin: Springer Verlag.
11. Beer, R. D. (1996). Toward the evolution of dynamical neural networks for minimally cognitive behavior. In P. Maes, M. Mataric, J. Meyer, J. Pollack and S. Wilson (Eds.), From animals to animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior (pp. 421–429). MIT Press.
12. Floreano D., & Mondada F. (1996). Evolution of homing navigation in a real mobile robot. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, (26:) 3: 396–407.
13. Slocum, A.C., Downey, D.C. & Beer, R.D. (2000). Further experiments in the evolution of minimally cognitive behavior: From perceiving affordances to selective attention. In J. Meyer, A. Berthoz, D. Floreano, H. Roitblat and S. Wilson (Eds.), From Animals to Animats 6: Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior (pp. 430–439). MIT Press.
14. Beer R. D. (2003). The Dynamics of Active Categorical Perception in an Evolved Model Agent. Adaptive Behavior, 11(4), 209–243.
15. Ziemke T., & Thieme M. (2002). Neuromodulation of Reactive Sensorimotor Mappings as a Short-Term Memory Mechanism in Delayed Response Tasks. Adaptive Behavior, 10(3–4), 185–199.
16. Hartland, C., Bredeche, N., & Sebag, M. (2009). Memory-enhanced Evolutionary Robotics: The Echo State Network Approach. In IEEE Congress on Evolutionary Computation, 2009. CEC '09 (pp. 2788–2795).
17. Greeff J. de, & Nolfi S. (2010). Evolution of Implicit and Explicit Communication in Mobile Robots. In Nolfi S. & Mirolli M. (Eds.), Evolution of Communication and Language in Embodied Agents (pp. 179–214). Springer Berlin Heidelberg.
18. Gigliotta O., Pezzulo G., & Nolfi S. (2011). Evolution of a predictive internal model in an embodied and situated agent. Theory in Biosciences, vol. 130(4), pp. 259–276 doi: 10.1007/s12064-011-0128-x PMID: 21604186
19. Chung J. R., & Choe Y. (2011). Emergence of Memory in Reactive Agents Equipped With Environmental Markers. IEEE Transactions on Autonomous Mental Development, 3(3), 257–271.
20. Ollion C., Pinville T., & Doncieux S. (2012). With a little help from selection pressures: evolution of memory in robot controllers. In Artificial Life (Vol. 13, pp. 407–414). MIT Press.
21. Lehman, J., & Mikkulainen, R. (2014). Overcoming deception in evolution of cognitive behaviours. Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation. New York, USA: ACM.
22. Wilson M. (2002). Six views on embodied cognition. Psychological Bulletin & Review, 9 (4): 625–636.
23. Maeda, N. (2012). External working memory and the amount of distributed cognition. In Proceedings of the 34th Annual Meeting of the Cognitive Science Society (pp. 1954–1959).
24. Risko E. F., & Dunn T. L. (2015). Storing information in-the-world: Metacognition and cognitive offloading in a short-term memory task. Consciousness and Cognition, 36, 61–74. doi: 10.1016/j.concog.2015.05.014 PMID: 26092219
25. Gilbert S.J. (2015). Strategic offloading of delayed intentions into the external environment. The Quarterly Journal of Experimental Psychology, 68(5), pp.971–992. doi: 10.1080/17470218.2014.972963 PMID: 25404057
26. Gilbert S.J. (2015). Strategic use of reminders: Influence of both domain-general and task-specific metacognitive confidence, independent of objective memory ability, Consciousness and Cognition, 33: 245–260. doi: 10.1016/j.concog.2015.01.006 PMID: 25666463

27. Antonelo, E., Schrauwen, B., & Stroobandt, D. (2008). Mobile robot control in the road sign problem using Reservoir Computing networks. In IEEE International Conference on Robotics and Automation, 2008. ICRA 2008 (pp. 911–916).

28. Bakker, B., Zhumatiy, V., Gruener, G., & Schmidhuber, J. (2003). A robot that reinforcement-learns to identify and memorize important previous observations. In 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings (Vol. 1, pp. 430–435 vol.1).

29. Linaker, F., & Jacobsson, H. (2001). Mobile Robot Learning of Delayed Response Tasks Through Event Extraction: A Solution to the Road Sign Problem and Beyond. In Proceedings of the 17th International Joint Conference on Artificial Intelligence—Volume 2 (pp. 777–782). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

30. Rylatt R. M., & Czarnecki C. A. (2000). Embedding Connectionist Autonomous Agents in Time: The "Road Sign Problem." Neural Processing Letters, 12(2), 145–158.

31. Duarte M., Oliveira S. M., & Christensen A. L. (2014). Evolution of Hybrid Robotic Controllers for Complex Tasks. Journal of Intelligent & Robotic Systems, 78(3–4), 463–484.

32. Blynel J., & Floreano D. (2003). Exploring the T-Maze: Evolving Learning-Like Robot Behaviors Using CTRNNs. In Cagnoni S., Johnson C. G., Cardalda J. J. R., Marchiori E., Corne D. W., Meyer J.-A., . . . Hart E. (Eds.), Applications of Evolutionary Computing (pp. 593–604). Springer Berlin Heidelberg.

33. Paine R. W., & Tani J. (2005). How Hierarchical Control Self-organizes in Artificial Adaptive Systems. Adaptive Behavior, 13(3), 211–225.

34. Bonani, M., Longchamp, V., Magnenat, S., Retornaz, P., Burnier, D., Roulet, G., et al. (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 (IROS) (pp. 4187–4193).

35. Massera G., Ferrauto T., Gigliotta O., & Nolfi S. (2013). FARSA: An Open Software Tool for Embodied Cognitive Science. In Advances in Artificial Life, ECAL 2013 (pp. 538–545). MIT Press.

36. Sperati V., Trianni V., & Nolfi S. (2008). Evolving coordinated group behaviour through maximization of mean mutual information. Special Issue on Swarm Robotics, Swarm Intelligence Journal, 4(2), 73–95.

37. Beer R. D., & Gallagher J. C. (1992). Evolving Dynamical Neural Networks for Adaptive Behavior. Adaptive Behavior, 1(1), 91–122.

38. Gigliotta O., & Nolfi S. (2008). On the Coupling Between Agent Internal and Agent/ Environmental Dynamics: Development of Spatial Representations in Evolving Autonomous Robots. Adaptive Behavior—Animals, Animats, Software Agents, Robots, Adaptive Systems, 16(2–3), 148–165.

39. Nasr, G. E., Badr, E. A., & Joun, C. (2002,). Cross Entropy Error Function in Neural Networks: Forecasting Gasoline Demand. In FLAIRS Conference (pp. 381–384). Chicago

40. Wood E. R., Dudchenko P. A., Robitsek R. J., & Eichenbaum H. (2000). Hippocampal Neurons Encode Information about Different Types of Memory Episodes Occurring in the Same Location. Neuron, 27 (3), 623–633. PMID: 11055443

41. Pfeiffer B. E., & Foster D. J. (2013). Hippocampal place cell sequences depict future paths to remembered goals. Nature, 497(7447), 74–79. doi: 10.1038/nature12112 PMID: 23594744

42. Wikenheiser A. M., & Redish A. D. (2015). Decoding the cognitive map: ensemble hippocampal sequences and decision making. Current Opinion in Neurobiology, 32, 8–15. doi: 10.1016/j.conb.2014.10.002 PMID: 25463559

43. Von Hofsten C., & Ronnqvist L. (1988). Preparation for grasping an object: A developmental study. Journal of Experimental Psychology: Human Perception and Performance, 14 (4): 610–261 PMID: 2974872

44. Jerde T. E., Soechting J. F., & Flanders M. (2003). Coarticulation in fluent fingerspelling. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 23(6), 2383–2393.

45. Tversky B. (2011). Visualizing Thought. Topics in Cognitive Science, 3(3), 499–535. doi: 10.1111/j.1756-8765.2010.01113.x PMID: 25164401