On the evolution of autonomous time-based decision-making and communication in collective robotics

by

Christos Ampatzis

Université Libre de Bruxelles
Faculté des Sciences Appliquées
IRIDIA
CP 194/6, Avenue Franklin Roosevelt 50, 1050 Brussels, Belgium
campatzi@ulb.ac.be

Supervised by

Professor Marco Dorigo

Directeur de Recherches du FNRS
Université Libre de Bruxelles
Faculté des Sciences Appliquées
IRIDIA
CP 194/6, Avenue Franklin Roosevelt 50, 1050 Brussels, Belgium
mdorigo@ulb.ac.be

Doctor Elio Tuci

ISTC-CNR
via San Martino della Battaglia 44, 00185 Roma, Italy
elio.tuci@istc.cnr.it

A thesis submitted in partial fulfilment of the requirements of the
Université Libre de Bruxelles, Faculté des Sciences Appliquées for the degree of

PhD in Sciences de l'Ingenieur

Abstract

In this thesis, we use evolutionary robotics techniques to automatically design and synthesise behaviour for groups of simulated and real robots. Our contribution will be on the design of non-trivial individual and collective behaviour; decisions about solitary or social behaviour will be temporal and they will be interdependent with communicative acts. In particular, we study time-based decision-making in a social context: how the experiences of robots unfold in time and how these experiences influence their interaction with the rest of the group. We propose three experiments based on non-trivial real-world cooperative scenarios. First, we study social cooperative categorisation; signalling and communication evolve in a task where the cooperation among robots is not *a priori* required. The communication and categorisation skills of the robots are co-evolved from scratch, and the emerging time-dependent individual and social behaviour are successfully tested on real robots. Second, we show on real hardware evidence of the success of evolved neuro-controllers when controlling two autonomous robots that have to grip each other (autonomously self-assemble). Our experiment constitutes the first fully evolved approach on such a task that requires sophisticated and fine sensory-motor coordination, and it highlights the minimal conditions to achieve assembly in autonomous robots by reducing the assumptions *a priori* made by the experimenter to a functional minimum. Third, we present the first work in the literature to deal with the design of homogeneous control mechanisms for morphologically heterogeneous robots, that is, robots that do not share the same hardware characteristics. We show how artificial evolution designs individual behaviours and communication protocols that allow the cooperation between robots of different types, by using dynamical neural networks that specialise on-line, depending on the nature of the morphology of each robot. The experiments briefly described above contribute to the advancement of the state of the art in evolving neuro-controllers for collective robotics both from an application-oriented, engineering point of view, as well as from a more theoretical point of view.

Acknowledgements

The work presented in this thesis was mainly supported by the ANTS project.[1] During the first months of my contract, I was supported by the SWARM-BOTS project,[2] which I joined towards its end. My research work can be placed at the intersection of the ANTS project and two projects funded by the European Commission: the SWARM-BOTS project and the ECAgents project.[3] Even if the three projects focused on different scientific issues, there was a clear and challenging overlapping area in which I managed to install myself. From the beginning I was fascinated by the perspective and the openness of all of them, so I think that discovering the synergies between their objectives was a good decision.

The person that actually allocated me to the aforementioned projects was professor Marco Dorigo. Professor Dorigo offered me a PhD student position in fall 2003 without knowing me and after a simple face-to-face meeting at IRIDIA, and I am still very grateful to him for this. I would like to thank him for the guidance and advice he gave me during all the years that he was my thesis supervisor. Also, I would like to thank him for being open-minded and listening to my ideas even if there might have been a priori disagreements. Finally, I should also acknowledge his support and valuable advice for my career.

The other person without whom this thesis would have never existed (at least in its present form) is my co-supervisor Dr. Elio Tuci. Dr. Tuci is a very passionate researcher with a pure drive about what he is doing and this greatly inspired me and influenced me as a person and as a researcher. Looking back at the countless arguments we had concerning work, I can only grin when I see the large amount of work we produced together.

At an intermediate level between supervisor, colleague and friend I will also place Dr. Vito Trianni because he helped me, both when I started my PhD, but also many many times during its course, to find my way in research. It was very beneficial to collaborate with him on many projects.

IRIDIA is a vibrant place, full of young and dynamic researchers coming from all over the world, and this is something that helped me in many ways: I came in contact with people from different backgrounds, I improved my English and French and I can claim (in my CV) experience at an international environment. IRIDIA might be the last professional environment where I make real friends, and for this reason it will always hold a special place in my heart. I want to thank the amazing people I have met during my PhD studies there for being good friends and colleagues. I will start chronologically: Shervin Nouyan, who initiated me to IRIDIA and who is still a close friend even if we never managed to collaborate; Roderich Groß, who taught me the value of details; Thomas Halva Labella, to whom I owe all I know about Linux; Max Manfrin, whom I annoyed a lot occupying

---

[1] An "Action de Recherche Concertée", funded by the Scientific Research Directorate of the French Community of Belgium.

[2] A project funded by the Future and Emerging Technologies Programme (IST-FET) of the European Commission, under grant IST-2000-31010. The project lasted 42 months and was completed successfully in 2005. See also `http://www.swarm-bots.org`.

[3] A project funded by the Future and Emerging Technologies Programme (IST-FET) of the European Commission, under grant IST-FET-1940. The project lasted 51 months and ended in March 2008. See also `http://www.ecagents.org`.

his computer; Bruno Marchal, with whom I am never bored to discuss; Mauro Birattari, a great scientist and a very interesting person; Carlotta Piscopo, with whom I hope to engage in a cooking competition one day; Anders Lyhne Christensen, the super-efficient; Rehan O' Grady, who always makes me ashamed of my Greek accent in English; Alexandre Campo, with whom I have spent countless hours discussing science; Prasanna Balaprakash, with his eternal smile; Marco Montes de Oca, whom I have exploited, always complaining about all my problems; Francisco Santos, a musical soul-mate, always assisting in wine-tasting and discovering obscure corners of my motherland and hopefully future co-author of articles; Carlo Pinciroli, with whom I might one day collaborate after all; Thomas Stützle; Tom Lenaerts; Paola Pelegrini; Federico Vicentini; Giovanni Pini; Alvaro Gutiérrez; Christophe Philemotte; Muriel Decreton.

I would also like to thank Dr. Stefano Nolfi who was always willing to discuss with me, mainly about my work, and was always happy to give me comments and suggestions, professor Hugues Bersini for the interesting discussions about science, Belgian politics but also music and art, and professor Jean-Louis Deneubourg.

Another fact for which I'm grateful to IRIDIA is that it introduced me to my first employers after my contract ended, Pierre Sener and Mohamed Ben Haddou of MENTIS Consulting. I was very happy to work with them even for a short while and I wish them all the best for the future.

Five years in a rainy land passed very very fast and happily due to some very important people in my life: Dimitrios Biliouris, Dimitrios Chatzis and Vasilis Valdramidis (Grieken Buiten!). But of course, the main responsible for keeping me here is Joke; it is mainly she that tolerated the wild mood swings of a stressed PhD student, and I am eternally thankful to her for that. Of course, I would also like to thank my family and to acknowledge moral (and occasionally financial...) support - thank you Antigoni, Mama, Papa and bossy uncle Christos.

I will close the acknowledgements by thanking the unsung heroes, the people that helped me go through this without knowing they did. I am referring to the artists whose music, films or books accompanied me during the PhD—an alternative "commité d'accompagnement" from Dorigo, Bersini, Deneubourg. I am talking about writers whose books kept me awake at night, directors who kept me in when my friends were out; I am talking about bands who stole a lot of my time when I had to listen to their CDs for pleasure or to review them, or when I had to drive for hours to go catch them live somewhere in Germany or in The Netherlands. For ecological reasons (printing extra pages means more trees cut) I will not mention them; they are far too many and the process of selecting among them and classifying them would be as time consuming as the writing of this very thesis has been.

“Blessed are the poor in spirit.” The Higher Lifer even has equivalents for spiritual poverty. As a man of science, he tries to keep himself unbiased by his interests and prejudices. But that’s not all. Ethical poverty of spirit entails taking no thought for the morrow, letting the dead bury their dead, losing one’s life to gain it. The Higher Life can make parodies of these renunciations. I know; for I made them and actually took credit to myself for having made them. You live continuously and responsibly only in the other, Higher world. In this, you detach yourself from your past; you refuse to commit yourself in the future; you renounce your own identity, except as a Higher Lifer, and become just the succession of your states. A more than Franciscan destitution. Which can be combined, however, with more than Napoleonic exultations in imperialism. I used to think I had no will to power. Now I perceive that I vented it on thoughts, rather than people. Conquering an unknown province of knowledge. Getting the better of a problem. Forcing ideas to associate or come apart. Bullying recalcitrant words to assume a certain pattern. All the fun of being a dictator without any risks and responsibilities.

Aldus Huxley, *“Eyeless In Gaza”*, 1932

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The history of robotics is arguably traceable back to ancient mythology. According to the legends and descriptions appearing in Homer's *Iliad*, the Greek god Hephaestus (god of fire and metal works) created mechanical servants in the form of tripods which could move around on their golden wheels as if they were self-propelled. Hephaestus not only constructed the equivalent of service rendering robots but even created intelligent and sensible maids that could speak—mythical entertainment robots. Hephaestus also made *"Thalos"*, a copper giant protecting the island of Crete from intruders. This idea does not differ that much from the *"golem"*, an animated being created entirely from inanimate matter according to Jewish tradition. This creature was supposed to serve its master, or according to the very well known myth, defend the Prague Jewish ghetto.

From then on, science fiction took over: From the robot obeying the three laws of robotics in Asimov's *I, Robot*, to the anthropomorphic robot *"Maria"* in Fritz Lang's *Metropolis* that manages to convince the workers to engage themselves into a rebellion; all the way to *"The Terminator"* at the end of the 20^{th} century.

But how far are we really from artifacts able and ready to loudly and proudly proclaim what Kraftwerk sung back in 1977: "We are the robots"? Most scientists would agree that we are very far from that stage. Already getting from the stiff and rigid industrial manipulators to robots trying to invent their own grammar (Steels, 1999), it took a long time. And probably to arrive to self-conscious robots it will take even more time. But after many decades of robotic research, things are starting to move; increasingly more complex robotic behaviour is exhibited.

Surprisingly, this does not necessarily mean that robots get increasingly more complex. Scientific research has drifted away from the paradigm of a monolithic omni-potent robot; instead, the approach of considering groups of robots cooperating and complementing each other's skills has become prominent. In order to design such systems, scientists took inspiration from examples of natural organisms, where simple entities do amazing things by working together. The nest construction of honey bees and termites are formidable examples.

In the near future we can expect the appearance of robots that interact amongst themselves and with humans; that operate in real-world environments, like our world, escaping the four walls of the robotic arenas in research labs; that sense the world through many and very advanced sensors; that can use a large number of sophisticated actuators; that take complex decisions as individuals, but also as part of a collective.

Individual and collective decision-making in a real-world scenario could endow autonomous agents with the ability to change their behaviour in response to the information gained through repeated interactions among themselves and with their environment. Decisions about solitary or social behaviour may be temporal, and they will be entangled with communicative acts. As the environment changes, robots should also *adapt* their behaviour. For example, they might need to interact with their environment in order to "realise" that cooperation is needed in order to achieve their goals. We believe that in order to obtain adaptive decision-making, the latter should be autonomous and a result of the agent-agent and agent-environment interaction; decisions should not be pre-defined by the experimenter.

Designing controllers able to tackle such complexity in so many levels is a non-trivial task. Failure to predict all things that might go wrong, or that could simply occur and happen to the robots, will lead to robots unable to adapt to changing and unpredictable environments. Machine learning techniques (Mitchell, 1997), and more specifically automatic design techniques, could be a solution to the aforementioned problem. In the BBC television documentary following his book (Dawkins, 1986), Richard Dawkins discusses with John Holland—the "father" of genetic algorithms (Holland, 1975)—about evolving machines using a process similar to, or more accurately, inspired by Darwinian evolution. Back in 1986, a technique like that was at its infancy; today, it is embodied into the design approach called *evolutionary robotics* (ER).

ER is a method for automatically synthesising robot controllers (Nolfi and Floreano, 2000). It is based on the use of artificial evolution to find sets of parameters for artificial neural networks that guide the robots to the accomplishment of their task. It is inspired by the Darwinian principle of selective reproduction of the fittest individual in a population, and it can exploit the richness of solutions offered by the dynamic robot-robot and robot-environment interactions, which may not be *a priori* evident to the experimenter (Dorigo and Colombetti, 1998; Nolfi and Floreano, 2000; Dorigo et al., 2004; Harvey et al., 2005).

In this thesis, we use ER techniques to automatically synthesise and design behaviour for groups of simulated and real robots. Our contribution will be on the design of non-trivial individual and collective behaviour; decisions about solitary or social behaviour will be temporal and they will be interdependent with communicative acts. In particular, we study time-based decision-making in a social context: how the experiences of robots unfold in time and how they influence their interaction with the rest of the group. The tasks we propose are based on non-trivial real-world cooperative scenarios.

1.1 Original contributions and related publications

Notice that the experiments described do not necessarily follow from each other, but they are rather independent. However they all share certain important methodological aspects. First, in all of them we use evolutionary robotics to design collective behaviour in a rather prejudice-free fashion. This means that the bias introduced by the choices of the experimenter will be kept to a functional minimum. Second, in all experiments we make use of similar dynamical neural networks that can grant the robots the possibility to make time-dependent decisions, that is, Continuous Time Recurrent Neural Networks (CTRNNs). Third, all experiments make use of the same robotic platform, that is the *s-bot*, a small autonomous robot with self-assembling capabilities that will be presented in detail in chapter 2, section 2.3.1. Controllers are evolved in a simulated environment and, for two of our experiments, they are downloaded and tested on the real *s-bot*. The simulated *s-bot* used varies according to the demands of each experiment.

From a more high level point of view, it must be stressed that the experiments are all about time-dependent decision-making in a social context. Thus, the robots engage in collective actions, mediated by different types of communication. These types will range from simple coordination through minimal sensors without the use of explicit and dedicated communication channels, to complex acoustic signalling interaction.

The contributions brought forth by this thesis can be summarised in the following four points:

First, we evolved signalling and communication in a task where the cooperation among robots was not a priori required. We used the evolutionary robotics approach in order to co-evolve the robots' communication and categorisation skills from scratch. This means that artificial evolution had to design at the same time (i) individual time-based decision-making mechanisms, and (ii) signalling mechanisms in robots that emit signals as well as reacting mechanisms in robot that receive those signals. Using this design methodology, we managed to unveil a structural relationship between categorisation and communication, with the former being a pre-requisite for the latter. This research work is covered in chapter 4 of this thesis and in the following publications in peer-reviewed international journals and international conference proceedings:

- **C. Ampatzis**, E. Tuci, V. Trianni, and M. Dorigo. Evolution of Signalling in a Group of Robots Controlled by Dynamic Neural Networks. In E. Sahin, W. M. Spears, and A. F. T. Winfield, editors, Second International Workshop on Swarm Robotics at SAB 2006, volume 4433 of Lecture Notes in Computer Science, pages 173-188. Springer-Verlag, Berlin, Germany, 2006.

- V. Trianni, **C. Ampatzis**, A. Christensen, E. Tuci, M. Dorigo, and S. Nolfi. From Solitary to Collective Behaviours: Decision Making and Cooperation. In Advances in Artificial Life, Proceedings of ECAL 2007, volume LNAI 4648 of Lecture Notes in Artificial Intelligence, pages 575-584. Springer-Verlag, Berlin, Germany, 2007.

- **C. Ampatzis**, E. Tuci, V. Trianni, M. Dorigo. Evolution of Signaling in a Multi-Robot System: Categorization and Communication. Adaptive Behavior, volume 16, number 1, pages 5-26, 2008.

Second, we have shown the portability of evolved neuro-controllers, and in particular evolved CTRNNs, on real robots for tasks where the integration over time of the perceptions of the robots is required. In these tasks, the type of "memory" required is very different from tasks where a particular event has to be remembered for some time by the robot. Instead, the robots are supposed to continuously interact with their environments (and teammates) in order to eventually make decisions. This contribution has been covered in chapter 4 of this thesis and in the following publications:

- **C. Ampatzis**, E. Tuci, V. Trianni, and M. Dorigo. Evolving Communicating Agents that Integrate Information over Time: A Real Robot Experiment. In E. G. Talbi, P. Liardet, P. Collet, C. Fonlupt, and E. Lutton, editors, 7^{th} International Conference on Artificial Evolution, EA'05, pages 1-12, Lille, France, October 2005. University of Lille.
- **C. Ampatzis**, E. Tuci, V. Trianni, and M. Dorigo. Evolution of Signalling in a Group of Robots Controlled by Dynamic Neural Networks. In E. Sahin, W. M. Spears, and A. F. T. Winfield, editors, Second International Workshop on Swarm Robotics at SAB 2006, volume 4433 of Lecture Notes in Computer Science, pages 173-188. Springer-Verlag, Berlin, Germany, 2006.
- **C. Ampatzis**, E. Tuci, V. Trianni, M. Dorigo. Evolution of Signaling in a Multi-Robot System: Categorization and Communication. Adaptive Behavior, volume 16, number 1, pages 5-26, 2008.

Third, similarly to the above, we show on real hardware evidence of the success of evolved neuro-controllers when controlling two autonomous robots that have to grip each other (autonomously self-assemble). Our experiment constitutes the first fully evolved approach on such a task that requires sophisticated and fine sensory-motor coordination. All actuators, including the gripper that is part of the connection mechanism are directly controlled by the activation of neurons of the CTRNN controller. Moreover, our approach highlights the minimal conditions to achieve assembly in autonomous robots by reducing the assumptions a priori made by the experimenter to a functional minimum. The research line is covered in chapter 5 of this thesis and has led to the writing of the following articles:

- E. Tuci, **C. Ampatzis**, V. Trianni, A.L. Christensen, M. Dorigo: Self-Assembly in Physical Autonomous Robots: the Evolutionary Robotics Approach. In S. Bullock, J. Noble, R. Watson and M.A. Bedau, editors, 11^{th} International Conference on the Simulation and Synthesis of Living Systems (Artificial Life XI), pages 616-624, The MIT Press, Cambridge, MA.
- **C. Ampatzis**, E. Tuci, V. Trianni, A. L. Christensen, M. Dorigo. Evolving Autonomous Self-assembly in Homogeneous Robots. Revised version submitted to the Artificial Life Journal.

Fourth, we present the first work in the literature to deal with the design of homogeneous control mechanisms for morphologically heterogeneous robots, that is robots that do not share the same hardware characteristics. The research on autonomous self-assembly mentioned above, showed that the same controller, when downloaded on robots with identical morphology, can exploit their history of interactions, noise in the sensors/actuators and potential differences among the robots hardware in order to cause their specialisation or allocation of distinct roles. In this work, we explore the same approach, that is, the same dynamical neural network is ported on a group of simulated robots; however, the goal this time is to study specialisation in a group of heterogeneous robots. In our experiment, this corresponds to robots having different sets of sensors at their disposal, but that are controlled by the same dynamic neuro-controller. By using artificial evolution to design individual behaviours and the communication protocols that allow the cooperation between robots of each type, we obtained dynamical neural networks that specialise on-line, depending on the nature of the morphology of each robot. This research work has been covered in chapter 6 of this thesis and published in the following peer-reviewed conference proceedings and international journals:

- E. Tuci, **C. Ampatzis**, F. Vicentini, and M. Dorigo. Operational Aspects of the Evolved Signalling Behaviour in a Group of Cooperating and Communicating Robots. In P. Vogt, Y. Sugita, E. Tuci, and C. Nehaniv, editors, Symbol Grounding and Beyond: Third International Workshop on the Emergence and Evolution of Linguistic Communication, EELC 2006, volume 4211 of Lecture Notes in Artificial Intelligence, pages 113-127. Springer-Verlag, Berlin, Germany, 2006.

- E. Tuci, **C. Ampatzis**, F. Vicentini, and M. Dorigo. Evolved Homogeneous Neuro-controllers for Robots with Different Sensory Capabilities: Coordinated Motion and Cooperation. In S. Nolfi, G. Baldassarre, R. Calabretta, J. Hallam, D. Marocco, J.-A. Meyer, O. Miglino, and D. Parisi, editors, From Animals to Animats 9: 9^{th} International Conference on Simulation of Adaptive Behavior, SAB 2006, volume 4095 of Lecture Notes in Artificial Intelligence, pages 679-690. Springer-Verlag, Berlin, Germany, 2006.

- E. Tuci, **C. Ampatzis**, F. Vicentini, M. Dorigo. Evolving Homogeneous Neuro-controllers for a Group of Heterogeneous Robots: Coordinated Motion, Cooperation, and Communication. Artificial Life, 14(2):157-178, 2008.

Furthermore, it should be noted that the research presented in this thesis was partly supported by the SWARM-BOTS project, and thus the author of this thesis has contributed to the following collective publications (peer-reviewed conference proceedings and a book chapter):

- M. Dorigo, E. Tuci, R. Gross, V. Trianni, T.H. Labella, S. Nouyan, **C. Ampatzis**, J-L Deneubourg, G. Baldassarre, S. Nolfi, F. Mondada, D. Floreano, L.M. Gambardella. The SWARM-BOTS Project. In Proceedings of the First International Workshop on Swarm Robotics–The 8^{th} International Conference on the Simulation of Adaptive Behavior (SAB'04), 13-17 July 2004, Los Angeles, CA, USA, Volume 3342, pages 31-44.

- M. Dorigo, E. Tuci, V. Trianni, R. Gross, S. Nouyan, **C. Ampatzis**, T. H. Labella, R. O'Grady, M. Bonani, F. Mondada. SWARM-BOT: Design and Implementation of Colonies of Self-assembling Robots. Computational Intelligence: Principles and Practice, Gary Y. Yen and David B. Fogel, editors, IEEE Computational Intelligence Society, NY, 2006, pages 103–135.

Finally, even though not covered in this thesis since the research work presented falls out of its scope, original research work concerning the evolution of neural mechanisms for an iterated discrimination and categorisation task has been performed and presented in the following peer-reviewed conference paper:

- E. Tuci, **C. Ampatzis**, and M. Dorigo. Evolving Neural Mechanisms for an Iterated Discrimination Task: A Robot Based Model. In M. S. Capcarrere, A. A. Freitas, P. J. Bentley, C. G. Johnson, and J. Timmis, editors, Advances in Artificial Life: 8^{th} European Conference, ECAL 2005, volume 3630 of Lecture Notes in Artificial Intelligence, pages 231-240. Springer-Verlag, Berlin, Germany, 2005.

1.2 Thesis layout

This thesis is organised into eight chapters. Chapter 2 is dedicated to the background of the research work presented in the thesis. Here, we briefly present the collective and swarm robotics research fields and we describe the three projects in which our experimental work is situated and from which it draws its inspiration and motivations. We also detail, in section 2.7, the properties of evolutionary robotics, the design methodology used in all experiments presented. In chapter 3, we review the related work in the evolution of time-dependent neural structures and the evolution of signalling capabilities and communication in collective robotics. References to the particular literature that is only relevant for a specific experiment will be given inside the chapter in which the experiment is discussed.

Chapter 4 details the results of a first series of experiments aimed to unveil the relationship between categorisation and communication, and to study the co-evolution of time-dependent decision-making and collective behaviour, also reporting on experiments performed on real robots. In chapter 5, we go on to discuss the results of a second series of experiments in which we successfully apply evolved dynamical neuro-controllers on a real-world task, that is, the physical connection between two real autonomous robots. Chapter 6 deals with the design of homogeneous controllers for morphologically heterogeneous simulated robots.

In chapter 7, we present ideas on future work, backed up by two preliminary experiments on the evolution of functional self-assembly and the evolution of more complex forms of signalling in communicating and cooperating robots. Finally, in chapter 8, we draw conclusions.

Chapter 2

Background

In this chapter we give the background of the research work presented in this thesis. We place our experiments in the area of collective robotics, which we briefly present in section 2.1, but our work is also relevant for a sub-domain of collective robotics, namely swarm robotics, which we present in section 2.2. In sections 2.3 and 2.4, we present two completed projects funded by the European Commission, SWARM-BOTS and ECAgents, and the ANTS project (funded by the French community of Belgium), since our research has been conducted in the context of these projects; more precisely, it has been conducted at the intersection of the three projects. The chapter closes with the presentation of the design methodology that is used throughout all experiments detailed in this thesis, that is, evolutionary robotics (ER) (see section 2.7).

2.1 Collective robotics

The field of collective robotics focuses on the study of robotic systems that are composed of a number of autonomous robots which act together in order to reach a common goal.

A multi-robot approach can have many advantages over a single-robot system. First, a monolithic robot that could accomplish various tasks in varying environmental conditions is difficult to design; it can in principle be cheaper to build several simpler robots that can complement each other in order to achieve a common goal than to build one monolithic robot that can do everything by itself. Moreover, the single-robot approach suffers from the problem that even small failures of the robotic unit may prevent the accomplishment of the whole task. On the contrary, in a group of robots there can be enough redundancy to avoid having one single point of failure. This means that, should one robot fail (due to, e.g., a mechanical fault), there can be other robots that can replace it or take over its duties. Clearly, a single robot system would completely fail if its functional unit experience technical difficulties. Finally, parallelism can lead to more efficient and faster task execution (V.Jones and Matarić, 2006); we can even have heterogeneous teams of robots working together by complementing each other (see Parker, 1998; Tang and Parker, 2005, for example).

However, as the size of the group increases, new challenges arise for the control program that guides the group to the accomplishment of its task. In particular, the controller has to take into account the interactions among robots. For example, the group might have to learn social rules to reduce possible negative interferences (e.g., due to overcrowding, see Matarić, 1997a; Labella et al., 2006, for examples on collective foraging). Moreover, should the control be centralised, then it is completely dependent on the communication system between the central controller and the agents; should the central controller or the communication system fail, the whole system would collapse. Finally, a system using complex communication protocols might not be able to scale well with the number of robots, due to, e.g., limited communication bandwidth.

To address such issues, robotics research began to take inspiration from biological systems. Nature abounds with examples of decentralised self-organising systems (Camazine et al., 2001), where animals exhibit complex behaviour at the collective level, regulated by simple and local communication. Probably the most typical example is the collective behaviour exhibited by natural *swarms* of fish schools, bird flocks and social insects as ants, termites, wasps and bees (see figure 2.1). Swarms in nature stand as fascinating examples of how collectively intelligent systems can be generated from a large number of simple individuals following simple and local rules. Such systems are not centralised and there are no hierarchical structures as leaders. Still, these systems can display very coherent and very complex behaviour. Inspiration from such systems gave rise to the domain of swarm robotics which we present in the following section.

2.2 Swarm robotics

Swarm robotics is a novel approach to the design and implementation of robotic systems. These systems are composed of *swarms* of robots which tightly interact and cooperate to reach their goal. Swarm robotics draws direct inspiration from biology and its foundations are firmly rooted in *swarm intelligence*, the discipline which aims at the design of multi-agent systems inspired by the efficiency and robustness observed in social insects, such as ants, bees, wasps and termites (Bonabeau et al., 1999). As an approach, swarm robotics emphasises aspects like decentralisation of control, local and limited communication among

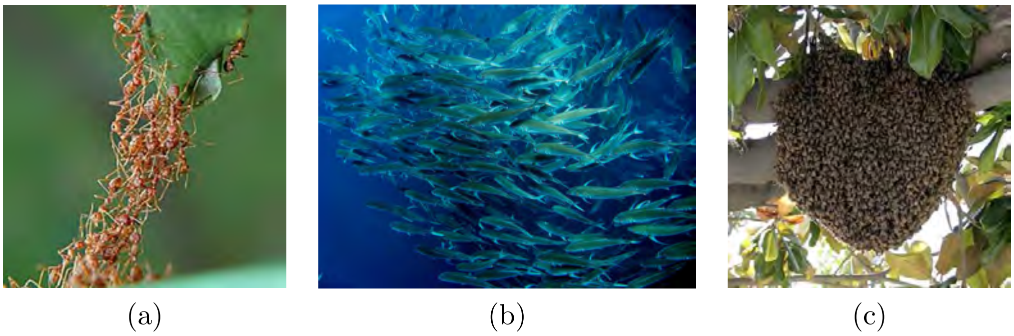

(a) (b) (c)

Figure 2.1: Natural swarms: (a) Ants forming a chain with their own bodies. (b) A fish school. (c) A swarm of bees.

robots and simplicity of the agents and the rules controlling the agents. The desired properties of the approach that derive from the aforementioned aspects are: (i) flexibility, that is, ability to adapt to new or changing environmental conditions, (ii) robustness, that is, ability to continue functioning even in case of failure of some system components, and (iii) scalability, that is, ability to gracefully handle growing amounts of agents. In a swarm robotic system, individual robots may be fully autonomous. Some tasks, however, may only be solvable by larger groups of robots that can overcome the physical constraints or limited abilities of individual robots.

Swarm robotics could be considered an instance of the more general field of collective robotics. Dorigo and Şahin (2004) define some criteria to which systems must comply in order to be considered swarm robotic systems. These criteria are: (i) scalability and control of large numbers of robots; (ii) homogeneous consistence of the groups; (iii) need for cooperation when carrying out certain tasks; (iv) use of local and limited sensing and communication abilities. As the authors highlight, these criteria cannot and should not be used as a checklist for determining whether a certain study falls into the category of swarm robotics research or not. Instead, they can give a flavour of the qualitatively distinctive features of swarm robotics systems, when compared to other multi-robot systems.

The SWARM-BOTS project[1] that we present in the next section contributed to the advance of the state of the art in swarm robotics.

2.3 The SWARM-BOTS project

The SWARM-BOTS project aimed to study new approaches to the design and implementation of self-organising and self-assembling artifacts through the development of a swarm robotic system called *swarm-bot*. The project was situated in the domain of collective and swarm robotics with a clear focus on the integration of self-assembly into the collective behaviour exhibited by the robots.

2.3.1 The *swarm-bot* and the *s-bot*

A *swarm-bot* is defined as an artifact composed of a swarm of *s-bots*, mobile robots with the ability to physically connect to/disconnect from each other. Also, they can connect to/disconnect from static objects. *S-bots* have simple sensors and motors and limited computational capabilities. Their connection mechanism is used to assemble into a *swarm-bot* able to solve problems that cannot be solved by a single *s-bot*. A *swarm-bot* is an assembly of two or more *s-bots*.

The *s-bots* are wheeled cylindrical robots with a 5.8 cm radius, equipped with many sensors useful for the perception of the surrounding environment and for proprioception, a differential drive system, and a gripper by which they can grasp various objects or another *s-bot* (see figure 2.2a). The traction system is composed of both tracks and wheels. The combination of tracks and wheels provides the *s-bot* with a differential drive motion,

---

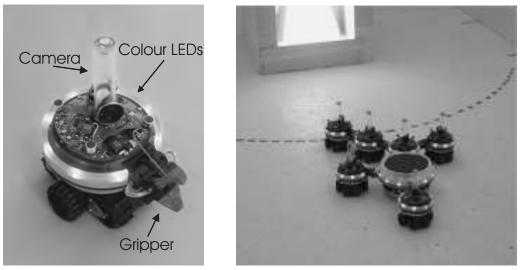
[1] A project funded by the Future and Emerging Technologies Programme (IST-FET) of the European Commission, under grant IST-2000-31010. The project lasted 42 months and was completed successfully in 2005. See also `http://www.swarm-bots.org`.

which is labelled *Differential Treels© Drive*. The treels are connected to the chassis, on which a cylindrical turret is mounted by means of a motorised joint. The turret can rotate with respect to the chassis, a property that allows the study of problems that were previously hard to tackle due to hardware constraints, as, for example, the coordinated motion of an assembled structure. The *s-bot* is equipped with a traction sensor that can detect the direction and intensity of the pulling force that the turret exerts on the chassis. The functionality of the rotating turret and the traction sensor were not used in our experimentation.

Fifteen infra-red proximity sensors are distributed around the rotating turret, and can be used to detect obstacles and/or other *s-bots*. Four proximity sensors (referred to as floor or ground sensors) are placed under the chassis and can be used for the perception of holes in the floor. These sensors can also detect the floor's colour. Furthermore, an *s-bot* is provided with eight light sensors distributed around the turret and an omni-directional camera mounted on the turret. The camera can perceive coloured blobs up to a distance of approximately 50 cm. On the body of the robot are also mounted four microphones and one loudspeaker useful for acoustic communication. The loudspeaker can emit a sound signal of variable frequency and intensity. The signal is perceived by the microphones and processed by the on-board CPU in order to discriminate the perceived frequency and intensity. Finally, temperature and humidity sensors, a 3-axis accelerometer and torque sensors on most joints are also available.

Mounted on the perimeter of the *s-bot*'s turret are eight groups of three LEDs that can light up in different colours (red, green and blue), allowing the robot to display coloured patterns useful for either communication with other robots or simply their perception by other robot's cameras (see figure 2.2a).

The rigid gripper is mounted on the turret and can be used for connecting to other *s-bots* or to some objects. The shape of the gripper matches the T-shaped ring placed around the *s-bot*'s turret, so that a firm connection can be established. The gripper can open and close its claws, but it also has a degree of freedom for lifting the grasped objects. Inside the gripper claws there is an optical barrier, which is a hardware component composed of two LEDs and a light sensor. The optical barrier can provide useful information about the



(a) (b)

Figure 2.2: (a) A picture of an *s-bot*, showing the gripper, the colour LEDs and the omni-directional camera mounted on the turret. (b) An assembled structure (i.e., a *swarm-bot*) transporting an object too heavy to be moved by a single *s-bot*.

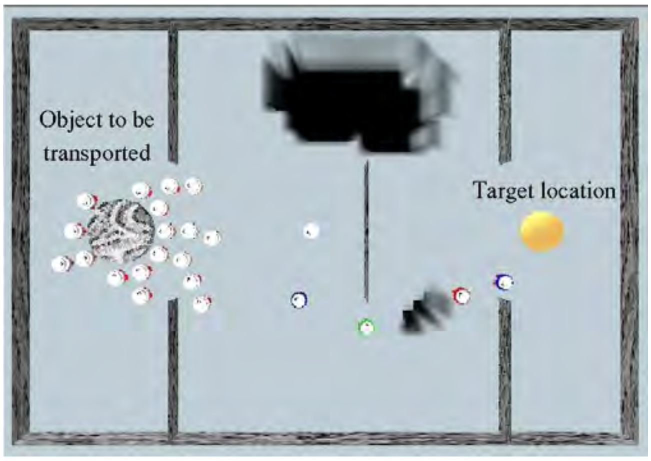


Figure 2.3: Picture of the scenario.

condition of the gripper and/or the presence of objects within the gripper claws. *S-bots* are also optionally provided with a flexible arm with three degrees of freedom, on which a second gripper is mounted—this actuator has not been considered for the experiments presented in this thesis.

The *s-bot* is a relatively small robot with a very rich sensory and actuator system that allows it to interact with its environment and other *s-bots* in many ways. The most innovative feature of the *s-bot* is the connection mechanism. The connection mechanism allows us to build controllers that will let the *s-bots* autonomously self-assemble. It also allows us to control the motion of assembled *s-bots* moving coordinately. A *swarm-bot* can solve tasks beyond the capabilities of the single *s-bots*. It can, for example, navigate on rough terrain, where a single robot would topple over, cross gaps too large for the size of the single robot, or transport objects too heavy to be moved by an *s-bot* (see figure 2.2b for a depiction). A *swarm-bot* can also reconfigure along the way when needed. For example, it might have to adopt different shapes in order to go through a narrow passage or overcome an obstacle. Finally, it should be mentioned that there might also be occasions in which a swarm of independent *s-bots* is more efficient: for example, when searching for a goal location or when tracking an optimal path to a goal.

2.3.2 Results and the final scenario

The project's empirical work focused on the study of the following cooperative tasks: division of labour (Labella et al., 2006), aggregation (Dorigo et al., 2004), coordinated motion (Dorigo et al., 2004; Baldassarre et al., 2007), autonomous self-assembly (Groß et al., 2006a), cooperative transport (Tuci et al., 2006), path formation (Nouyan et al., 2008a) and cooperative hole avoidance (Trianni et al., 2006; Trianni and Dorigo, 2006). The final goal of the project was accomplished with the successful realisation of a scenario that is depicted in figure 2.3 and that can be summarised as follows:

> A swarm of up to 35 *s-bots* must transport a heavy object (called the prey)

> from its initial location to a goal location (called the nest). There are several possible paths from the initial to the goal location and these paths may have different lengths and may require avoiding obstacles and holes. The weight of the object is such that its transportation requires the coordinated work of at least n *s-bots*, where n is a parameter.

The above scenario was solved by 12 real *s-bots* (see Nouyan et al., 2006, 2008b, for details, and figure 2.4 for snapshots describing the solution) and it included the transport of a prey object to a nest location. The robots form chains that find the prey in the environment and subsequently they connected to it and transport it to the nest location.

(a) (b) (c)

Figure 2.4: The scenario solved by 12 *s-bots* (see Nouyan et al., 2006, 2008b, for details). (a) The initial setup: the object to be transported (prey) is not connected to the nest; (b) a chain that connects the prey and the nest is constructed; (c) the robots after connecting to the prey, transport it and finally bring it to the nest. The images are snapshots taken from a video complementing (Nouyan et al., 2006).

2.4 The ECAgents project

The ECAgents project[2] aimed at: (i) developing a new generation of agents able to evolve autonomously, self-organise, and operate reliably in dynamic environments; (ii) setting up the conditions that allow a population of agents to develop a shared communication language and to share knowledge; (iii) identifying new methods and algorithms that allow to engineer systems able to self-organise and to display properties emerging from the interactions between themselves and with the external environment.

In order to study the role of communication, the project considered (i) agents involved in collective tasks; (ii) communication systems that are not pre-designed by the experimenter; instead they emerge from the interactions of the agents among themselves and with the external environment; (iii) self-organising and evolving communication conventions and underlying ontologies; (iv) experimentation on real agents (e.g., robots).

---

[2] A project funded by the Future and Emerging Technologies Programme (IST-FET) of the European Commission, under grant IST-FET-1940. The project lasted 51 months and ended in March 2008. See also `http://www.ecagents.org`.

Unlike the SWARM-BOTS project, the ECAgents project was not confined in the domain of collective (or swarm) robotics; it rather tackled more general scientific questions related to communication among agents, where agents could be humans, animals, robots, software agents or other technological artifacts.

On a more conceptual axis, it is important to notice that the experiments and the studies performed during the course of the project contribute to better modelling the evolution of communication and language in natural organisms. In particular, they help to shed light on open questions concerning the relationships between, on the one hand, the evolution and the development of perceptual, cognitive and motor capabilities and, on the other hand, the emergence of a communicative system and possibly language in a population of agents. Given that it is impossible to reconstruct the evolutionary route to communication and language, these experiments on artificial agents can be valuable and significant tools at our disposal to form plausible hypotheses about the origin of language.

> Modelling simple pre-linguistic communication is useful because it puts further constraints on theories of how language itself evolved—as things stand there is no room for far too many plausible possibilities. Mathematical and simulation modelling are necessary steps if we are to go beyond an impasse in which the proponents of competing theories merely trade rhetoric.
>
> (Noble, 2000a)

2.4.1 Major project outcomes

The project addressed many diverse but also inter-connected scientific questions. In the following, we mention some of the major project outcomes:[3]

- The study of chemical communication which plays a central role in the transition from solitary to social behaviour (see Millor et al., 2006, for example).
- The study of the role of communication network topologies in multi-agent systems playing language games (see Dall'Asta et al., 2006, for example).
- The identification of selective scenarios for the origin of language (see Szamado and Szathmary, 2006, for example).
- The study of language and opinion dynamics in multi-agent systems through statistical physics and complex network theory (see Baronchelli et al., 2006, for example).
- The identification of prerequisites for observing the emergence of communication in a group of initially non-communicating agents (see Mirolli and Parisi, 2008; Floreano et al., 2007; Ampatzis et al., 2008, for examples), and for the evolution of a communication system with the characteristics of human language (see Steels, 2007, for example).

---

[3] Notice that this list is non-exhaustive; we encourage the interested reader to visit `http://www.ecagents.org` for more information.

- The co-evolution of behavioural and communicative skills from scratch (see Nolfi, 2005; Trianni and Dorigo, 2006; Ampatzis et al., 2008, for examples).
- The study of how a stable and reliable communication system can evolve from scratch despite possible conflicts of interest between individuals (see Floreano et al., 2007; Mirolli and Parisi, 2008, for examples).

2.5 The ANTS project

The ANTS project has as a goal to study and to establish the foundations of ant algorithms. Ant algorithms are inspired by the collective behaviour of social insects and they belong to the wider research field of swarm intelligence (Bonabeau et al., 1999; Garnier et al., 2007). They are targeted on engineering applications and they are applied to real-world optimisation problems (see Dorigo and Stützle, 2004, for details).

The objectives of the project are to:

- Study the theoretical properties of already established algorithms such as the ant colony optimisation (Dorigo and Stützle, 2004).
- The extensive evaluation of ant algorithms that are not yet sufficiently applied and tested on concrete problems.
- The identification and study of new ant algorithms.

Apart from the goals mentioned above, another goal of the project is to further the understanding of self-organisation in social insects (Camazine et al., 2001) in order to apply the principles extracted to engineering problems. Such problems include: NP-hard optimisation problems, the decentralised control of a swarm of robots and dynamic task allocation of resources.

A better understanding of the principles underlying collective behaviour in social insects (and non-human animals in general) could inspire engineers trying to design intelligent robotics systems. As stated in Arkin (1998), *"The study of animal behaviour can provide models that a roboticist can operationalise within a robotic system"* (see Nouyan et al., 2008a, for an example). Thus, the study of animal behaviour, and for this particular project social insect group behaviour, can offer a lot to collective and swarm robotics.

2.6 At the intersection of the ANTS, the SWARM-BOTS, and the ECAgents projects

Even if the aforementioned projects had different goals and activities, there is a clear domain of overlap, which is related to the communication between robots in a collective robotics scenario. Communication is an issue of central importance in collective robotics, since it allows and regulates the switch from solitary to social behaviour. This domain lies at the intersection of the three projects and one can imagine a lot of synergy between the research lines. This intersection is exactly where the research work presented in this thesis is situated.

In collective robotics research, the coordination of the activities in a group of robots requires the definition of signalling and communicative behaviour among the individuals. Looking at social insects and at the complexity they achieve even without the use of global, direct or explicit communication, we can imagine that the communication strategies in a robotic group need not be particularly complex; simple forms of communication or no explicit communication at all can be enough to obtain the coordination of the activities of the group (Kube and Zhang, 1997). This is the case for swarm robotics, which focuses on simple and local communication, that can scale up with the number of agents involved. Scalability of the communication protocol is essential to achieve scalability of the system as a whole (one of the desirable properties of swarm robotic systems).

The two IST-FET projects have shared the same methodology to pursue their objectives, that is, artificial evolution. Evolutionary robotics has been used in the context of the SWARM-BOTS project from an engineering perspective, that is, to design controllers for groups or swarms of cooperating robots (see Tuci et al., 2006; Trianni et al., 2006; Dorigo et al., 2004; Baldassarre et al., 2007; Ampatzis et al., 2005, for examples). In the ECAgents project, ER has been employed as a methodology to investigate hypotheses on the evolution of communication in a group of agents (see Marocco and Nolfi, 2006a; Floreano et al., 2007; Mirolli and Parisi, 2008; Ampatzis et al., 2008, for examples). However, in many cases, even if the stress is on evolving communication, the resulting emergent signalling phenomena also have an engineering value as they can be used to regulate the flow of information within a group of cooperating robots.

The two projects have also shared resources in order to accomplish their goals. The *s-bot*, the mobile autonomous robot produced in the course of the SWARM-BOTS project, has been extensively used in the experiments performed in the course of the ECAgents project.

Furthermore, the two IST-FET projects share the same interest in designing self-organising robotic systems with the ANTS project. Camazine et al. (2001) define self-organisation as *"a process in which pattern at the global level of a system emerges solely from numerous interactions among the lower-level components of the system"*. Self-organisation is observed in many animal societies (as, for example, social insects like ants, bees or termites). From an engineering perspective, there are many advantages in designing a decentralised self-organising robotic system, as, for example, ability to adapt to varying environmental conditions and robustness to individual failure (Dorigo et al., 2004; Trianni and Dorigo, 2006).

In this thesis, as we have already said previously, ER will be the design methodology employed. Also, we will make use of the real *s-bot*, or simulated versions of it, as the robotic platform on which to test our design methodology and the obtained controllers. All the experiments that will be detailed in the following chapters of this thesis will treat the design of controllers for autonomous groups of robots, thus they will be experiments contributing to the advance of research in collective robotics. The design of the experiments will draw inspiration from the observation of social insect societies; in particular, we will employ simple communication mechanisms to obtain collective decision-making and coordination in a group of robots. The niche of these experiments lies exactly at the intersection of the two IST-FET projects (SWARM-BOTS and ECAgents) with the ANTS project.

2.7 Evolutionary robotics as a design methodology

In this section we present the design methodology that is used throughout all the experiments detailed in this thesis, that is, evolutionary robotics (ER). In section 2.7.1 we briefly describe the ER methodology, in section 2.7.2 we present its operating principles and its properties, and in section 2.7.3 we justify why it was selected as the design method utilised in our research work, by highlighting its differences to other approaches. Notice that we will not provide the reader with a full literature review of ER research; instead, in chapter 3 we will cover those research works which are relevant to the research presented in this thesis.

2.7.1 What is evolutionary robotics?

The problem of defining a controller for a robotic system has been approached from many different directions: motion planners (see Choset et al., 2005; LaValle, 2006, for overviews), artificial potential control (see Spears et al., 2004, for an example), behaviour-based robotics (see Brooks, 1986, 1991; Arkin, 1998; Balch and Arkin, 1998) and reinforcement learning systems (Kaelbling et al., 1996; Matarić, 1997b; Dorigo and Colombetti, 1998; Sutton and Barto, 1988) are only some examples of the possible ways of controlling a robot. Among these, evolutionary robotics (ER) is a technique for the synthesis of robot controllers; ER is a methodological tool to automate the design of robots' controllers (Nolfi and Floreano, 2000). It is based on the use of artificial evolution to find sets of parameters for randomly generated sets of controllers. The controllers may consist of rule sets, or can be decision trees, but the most commonly used are artificial neural networks (ANNs) due to their versatility, generalisation capabilities and tolerance to noisy sensory input. So, if the controller is represented as an ANN, an evolutionary algorithm can be applied in order to optimise the weights, and possibly the morphology, of the networks that guide the robots to the accomplishment of their task.

ER is inspired by the Darwinian principle of selective reproduction of the fittest individual in a population. This means that the individual that adapts best to its environment has more chances to reproduce and to pass its genetic material to subsequent generations. The process of searching the design space by mimicking natural evolution is generally referred to as evolutionary algorithm (Holland, 1975; Goldberg, 1989). In figure 2.5 we schematically describe the way artificial evolution works. In particular, we describe how a population of n genotypes, each encoding the control system (and sometimes the morphology) of the robots is generated by the population of the previous generation. Each genotype of the population at generation k ($GEN\ k$) is used to build an artificial neural network (ANN), that will control each robot (this is the so called homogeneous approach, where each robot shares the same controller). The robots are tested on a certain task, their performance is evaluated, and a fitness score (F) is attributed to the genotype. This procedure is completed independently for all n genotypes, until we have a fitness score

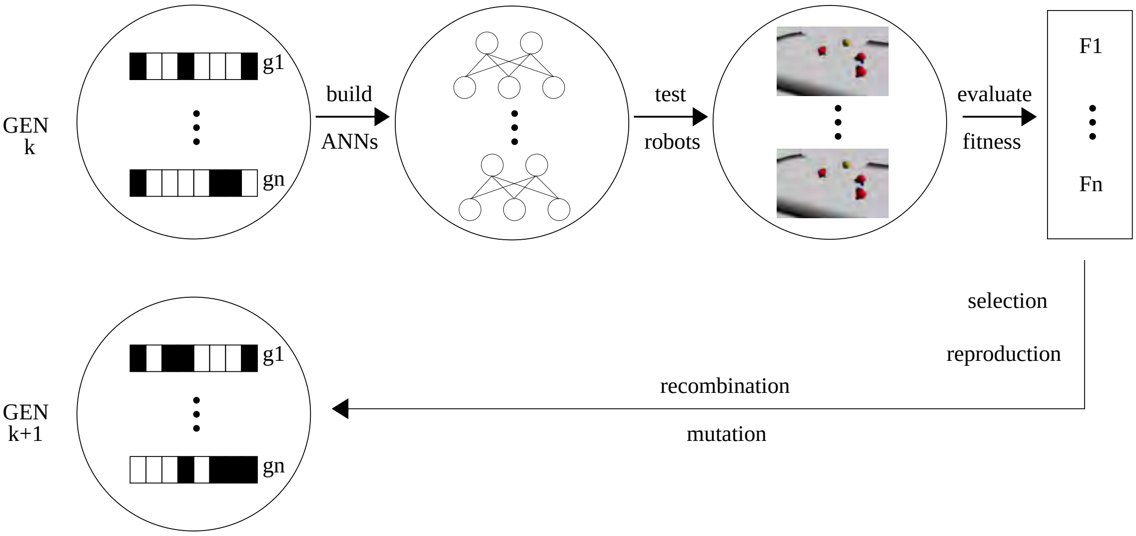


Figure 2.5: Depiction of how artificial evolution typically works. The picture shows how a population of controllers is generated from the previous generation.

attributed to every individual controller. Subsequently, selective reproduction is applied (fittest controllers reproduce with a higher frequency), and controllers generate copies of their genetic material, which are modified by genetic operators, such as mutation and crossover. In this way, we arrive to the next generation of genotypes ($GEN\ k+1$). Artificial evolution usually starts with a random population of individuals and terminates when a satisfying controller is found that meets the requirements stated by the experimenter in the performance evaluation.

2.7.2 Properties of evolutionary robotics

This section is dedicated to the presentation of three very important properties of ER: autonomous decision-making (production of solutions little or not biased by the experimenter) in section 2.7.2.1, co-evolution of communicative and non-communicative behaviour in section 2.7.2.2, and efficient solution of the design problem in section 2.7.2.3. The above properties are in fact the reasons why we chose ER as the design methodology to pursue our experiments on the design of time-based decision-making and communication.

2.7.2.1 Evolutionary robotics and autonomous decision-making

With respect to other design methods (e.g., back-propagation supervised learning), ER provides the methodological tools to generate control structure for artificial agents such as autonomous robots in a relatively prejudice-free fashion (Harvey et al., 2005). For example, ER does not require the designer to make strong assumptions concerning what behavioural and/or communication mechanisms are needed by the robots. The experimenter defines the characteristics of a solitary/social context in which robots are required to cooperate. The agents' mechanisms for solitary and social behaviour are determined by an evolutionary process that favours (through selection) those solutions which improve an agent's or group's ability to accomplish its task (i.e., the fitness measure).

> The use of artificial evolution minimises the incorporation of design prejudices and constraints, as the subtleties and tweaking of architectural detail are left to the blind forces of evolution, guided only by the selection constraints imposed by the experimenter on the behaviour, not on the mechanism.
>
> (Harvey et al., 2005)

The above properties give rise to a new realm of possibilities: the automatic process can find solutions to the problem at hand that were not a priori evident to the experimenter (Nolfi and Floreano, 2000; Dorigo et al., 2004). This is due to the fact that the process can exploit the richness of interactions among robots of a group and between robots and the environment (Trianni and Dorigo, 2006).

It should be noted that also other reinforcement learning techniques (RL), different than evolutionary computation, such as those presented in (Sutton and Barto, 1988, e.g., Q-learning), only need an evaluation of how good or bad the robots are doing at a given time; in case a neural network controller is used, the learning algorithm is based on a global evaluation of the network response, not on a definition of the exact output (Nolfi and Floreano, 2000). Thus these techniques also require limited supervision. However, ER has the advantage that, in principle, there is no constraint on what could be learnt; for example, robot shape, hardware specifications and neural network type, size and connectivity could be subject to evolution as well.

Of course, in order for the evolutionary process to be prejudice-free, or unbiased by the experimenter, certain conditions must hold. In ER, the experimenter must come up with a fitness function which will guide the evolutionary process. The design of the fitness function will severely influence the resulting control structures and so it can be considered an unavoidable bias. For example, if the designer of the fitness function encodes in it the communication protocols the robots should use, then it is evident that the designer demands that these protocols appear in the resulting controllers. Thus, this could be considered equivalent or similar to "hard-coding" these communication rules into the robots' control structures. On the contrary, we can imagine fitness functions that do not interfere with the individual and social skills the robots can develop in order to solve the task at hand. Such functions could guide the evolutionary process to areas of the fitness landscape that correspond to solutions of the task, but without interfering with *how* the task should be solved. In that case, more freedom is left to the automatic process.

Furthermore, as we will see in the forthcoming chapter, the choice of the artificial neural network that will guide the robots in the accomplishment of their task has a significant influence on the type of solution the automatic design process produces. For example, using a perceptron as control structure, excludes the possibility that the robots can perform time-dependent decision-making, because the outputs of a perceptron only depend on its current inputs. Thus, the universe of possible solutions artificial evolution will search into is significantly narrowed down. Of course, we should stress again that also the network's morphology can be evolved, possibly increasing even more the autonomy of the automatic process, but this approach has not been followed in the experiments presented in this thesis.

Finally, we would like to stress that less designer involvement in a system's functioning might result in agents (and systems) with increased autonomy (Boden, 2008).[4] Of course, as stated in (Di Paolo and Iizuka, 2008), *"current work in autonomous robotics based on ideas of automatic synthesis of design (e.g., evolutionary robotics) and dynamical systems approaches to cognition, is still far from achieving or even modelling autonomy in the strong sense"*. This is particularly true when robots are confronted with goal-oriented tasks, as is the case in practically all engineering robotics applications. However, the authors go on suggesting that ER methods might be *"the surest route to this goal"* (achieving autonomy).

2.7.2.2 Co-evolution of communicative and non-communicative behaviour

A very important property of the ER approach is that it permits the co-evolution of communicative and non-communicative behaviour; different strategies can co-adapt because selection depends only on an overall evaluation of the group (Nolfi, 2005).

By using ER techniques, an experimenter interested in evolving both individual non-communicative behaviour and social communicative behaviour, might ask a general question as how a group of agents that have to solve a given task might develop forms of communication that enhance their adaptive capability. In fact, social behaviour and communication can be seen as adaptive skills, part of the behavioural repertoire of the agents, in close relation and interdependent with the solitary, non-communicative behaviour. As stated in (Nolfi, 2005):

> Communication and language can be properly understood by taking into account their relation with other important behavioural, social, and cognitive processes... Only by co-adapting their behavioural non-communicative and communicative abilities, individuals might develop a really useful communication system grounded in the physical and behavioural characteristics of communicating individuals and able to exploit active perceptual capabilities.

This co-evolution of behavioural skills paves the way for understanding how we arrive to communicative acts in cooperative groups of robots. This can be an important step to better understand the value of communication in a particular setup but also the context into which it takes place. Moreover, by co-evolving behavioural skills both at the solitary and the social level, we can obtain agents displaying complex behaviour at both levels. For example, certain individual cognitive mechanisms can serve as the backbone for the development of communication mechanisms; the latter in turn can lead to the further development of the individual mechanisms, and so on.

---

[4] The definition of autonomy in robotics is a very controversial issue and a difficult endeavour and it falls out of the scope of this thesis.

2.7.2.3 Solving the design problem

Finally, ER represents an effective solution to the design problem. This problem concerns the decomposition of the group's activities into interactions between individual agents and interactions of individual agents with the environment. Both types of interaction are dynamic, and therefore difficult or even impossible to predict for an external observer. In such a context, we believe that evolutionary robotics is the methodology to be exploited (Nolfi and Floreano, 2000; Harvey et al., 2005). ER bypasses the problem of decomposition at both the levels of finding the mechanisms that lead to the emergent global behaviour, and of implementing those mechanisms in a controller for the robots. In fact, ER relies on the evaluation of the system as a whole, that is, on the emergence of the desired global behaviour starting from the definition of the individual ones. More specifically, ER optimises the individual controller (for example the weights of an ANN) that will be cloned on all the robots (in case of the homogeneous approach), on the basis of the global behaviour which is the result of robot-robot and robot-environment interactions.

Thus, we can say that evolutionary robotics experiments can be designed adopting a "holistic" approach resulting in artificial systems that resemble natural self-organising systems. A property of the latter systems is that their functioning cannot be explained as being due to interactions between independently definable sub-parts; each part is (to some extent) dependent on other parts for its existence and for its identity and role in the system (Boden, 2008).

2.7.3 Discussion: advantages, limitations and qualities of ER

Evolutionary robotics provides us with an opportunity to couple an agent's dynamical system with the environment's dynamical system, through sensory-motor interactions. By exhibiting both situatedness and embodiment (Brooks, 1991), it evaluates a solution based on the agent's interaction with its environment (Harvey et al., 2005). Situatedness refers to robots that are perceiving the world through their sensors, rendering abstract representations useless since all they need in order to be in the world and display behaviour can be perceived. Embodiment on the other hand, refers to robots acting the world in which they are situated; by modifying it they introduce a feedback which they subsequently receive.

However, situatedness and embodiment are important concepts also for behaviour-based robotics, another widely used bottom-up method to design robot controllers (Brooks, 1991). We believe that the major difference between this approach and ER lies in the fact that ER can produce solutions which are less biased by the experimenter. This can potentially lead to more adaptive systems and to solutions to the problem at hand that were not a priori evident to the experimenter (see section 2.7.2.1).

In section 2.7.2.1 we also mention reinforcement learning techniques. These techniques also rely on an evaluation of how well or badly a robot is doing at a particular time, and therefore represent another approach to the bias-free creation of robotic control systems. Dorigo and Colombetti (1998) presented an experiment where step-by-step reinforcement learning helps a real robot perform phototaxis. The authors highlight that applying the same reinforcement learning algorithm to multiple robots performing phototaxis should be straightforward, without having to take into account the specifications of each agent and without having to directly write a control "program". That would indeed be true for a task where the robots are very little or not engaged in collective behaviour. However, once the degree of inter-dependence of behaviour starts increasing and the utility of the action of each robot is dependent upon the current and past actions of the other team members, applying reinforcement learning is not easy, because credit assignment is hard to define (Parker, 2002). Apart from credit assignment, another important challenge that arises when applying RL to a multi-robot domain is the prohibitively large state space (Matarić, 1997b; Fernández et al., 2005). To be able to deal with this issue, researchers define the space at a higher level of description, coming up with basic behaviours and conditions that trigger them (see Matarić, 1997a; Liu et al., 2005, for examples), or they limit it to use only attributes that the designer considers necessary and by defining a set of discretised actions (Fernández et al., 2005).

On the contrary, ER methods do not require behavioural decomposition: the experimenter does not have to come up with states or basic modules of behaviour and transition conditions between the modules. Relying on the evaluation of the group's behaviour only at the end, managing this way to bypass the design problem (see section 2.7.2.3), we believe that ER provides a framework for a less designer-biased automatic synthesis of controllers for multi-robot systems.

Concerning the engineering value of ER, we believe that the method has large potentialities that derive from its properties, discussed above. The experiments detailed in this thesis contribute to the advancement of the state of the art in the engineering applications of ER. However, in order for ER to be considered a viable alternative to manual design when it comes to real-world engineering applications, several challenges which Matarić and Cliff (1996) identified more than a decade ago, have still to be overcome. First, the ER approach still suffers from a lack of a substantial body of works reporting on experiments on real robots. Since performing evolution directly on the hardware (see Mondada and Floreano, 1995, for example) can be extremely time-consuming and even dangerous for the robot, researchers evolve their controllers in a simulated environment. Unfortunately, the majority of ER research works do not proceed to test these controllers on real robots, and they only report on experiments in simulation. The problem with this is that being successful in simulation does not guarantee that the evolved controllers can transfer to reality. This is because it is difficult to accurately simulate physical systems (Brooks, 1991) and also because possible inaccuracies and abstractions made in simulation can be exploited by the genetic algorithm; this in turn may result in systems that do not transfer to reality (Matarić and Cliff, 1996). Moreover, properly setting the noise levels in simulation based on samples from real sensors has also been identified as a pre-requisite for successful porting to reality (Miglino et al., 1995; Jakobi, 1997). However, this is not always an easy endeavour, as it heavily depends on the specific robotic hardware used (Matarić and Cliff, 1996).

Second, it is rather true that the complexity of the tasks solved by groups of robots controlled by evolved neuro-controllers is lower than the complexity achieved by other methods using modular or hand-coded controllers. In other words, what Matarić and Cliff

(1996) was claiming more than a decade ago, still holds:

> One of the main goals of the work on evolutionary robotics is to provide a methodology for automatically synthesising more complex behaviours than those that can be designed by hand. However, a survey of the results in the field to date does not show any demonstrations that have reached that goal.

Third, even if ER could in principle reduce the human effort required to design controllers, this is usually not the case, as *"the behaviours produced by current evolved controllers are, on the whole, relatively simple, and could have been designed by hand with the same or lesser amount of effort"* (Matarić and Cliff, 1996). In other words, the complexity achieved by ER approaches seems incommensurate with the effort expended in designing or configuring the evolutionary system. We believe that the reason for this lies in the nature of the involvement of the experimenter in the design process. While in other approaches, the experimenter would have to hard-code a functional solution to the problem, in ER the involvement is on the one hand reduced, but on the other hand it has to be more "intelligent". The experimenter has to identify: (i) the relevant aspects that have to be modelled, (ii) the sort of fitness function that can lead to efficient solutions, ideally without interfering with the mechanisms underpinning behaviour, and (iii) initial conditions that assure evolvability, i.e., the possibility to progressively synthesise better solutions starting from scratch. This latter point is not trivial and requires a lot of experience on behalf of the experimenter. Usually the processes described above include a lot of preliminary evolutionary simulations that require long computation time.

> Contrary to the uninformed perceptions at the time when ER was born, one cannot treat artificial evolution as a magic box capable of solving any problem one poses to it (and all one must do is just wait). Fortunately, failures to evolve a desired behaviour, if followed by some analysis of the behaviours that do evolve, often lead to a revelation of what are the problems one must overcome as a designer of an evolutionary regime.
>
> (Di Paolo and Iizuka, 2008)

The process of setting up the evolutionary experiment involves a deeper understanding of the nature of the task the robots face on behalf of the experimenter, who is not simply demanded to come up with *a* solution to this task. This might seem to hinder the applicability of the method on pure engineering tasks where finding one functional solution to the problem might suffice, but it allows for the research on more abstract questions. Some sample questions could be: What type of communication is necessary for the robots to accomplish a given task? What are the minimal conditions for coordination among robots with a given robotic hardware in a certain environment? What type of memory is necessary for the robots to solve a certain task?

Such questions are not only relevant in an engineering context; they could be equally relevant when roboticists are trying to answer questions related to biology, and in particular animal behaviour. This is because *"robots can be used as models of specific animal systems to test hypotheses regarding the control of behaviour"* (Webb, 2000). There are several examples in the literature where roboticists are moving in this direction; we will simply mention two works studying cooperative transport in ants and robots. Kube and Bonabeau (2000) demonstrated the sufficiency of minimal assumptions to reproduce ant capabilities such as repositioning dependent on prey size, with a robot model. The authors note that no formal description of the biological phenomenon had been developed. Similarly, Groß and Dorigo (2008a) arrived to the confirmation of the plausibility of a hypothesis that has not been further investigated by biologists, that *"group transport in social insects has evolved from situations in which solitary transporters, without being aware of each other, cooperatively transported a common prey."*

We believe that robotics research can contribute to the understanding of biological phenomena, regardless of the method employed to define robots' controllers. Still, we believe that ER can be particularly suitable for testing hypotheses concerning the nature of the underlying mechanisms that underpin the agents' behaviour. Our view is that ER models can be complementary to other analytical modelling tools at the disposal of biologists (e.g., game theory models, see Maynard-Smith, 1982, for example). The latter models allow biologists to predict the outcome of, for example, coordination/anti-coordination problems (Lewis, 1969), given the set of behavioural strategies available to the agent and the payoff corresponding to all the possible combinations of actions among the actors.

ER models have been recently used to predict the evolutionary conditions related to the emergence of communication in robots (Lipson, 2007; Floreano et al., 2007). This demonstrates exactly the amount of possibilities for synergy that exist between biology and robotics, especially when the evolutionary history of the research subject can hardly be reconstructed (e.g., communication and language). However, we cannot go as far as to claim that ER practitioners have in their hands a tool that is the artificial counterpart of natural selection. Artificial evolution is extremely simplified with respect to natural evolution and it would be naïve to claim that robots undergoing artificial evolution of their control structures experience the same environmental conditions (and dangers) as (simple) animals that live in the real world. This is especially true when in the majority of ER experiments the fitness reward goes directly to the group; unlike animals subject to natural selection (Dawkins, 1976), individual robots have no conflict of interest.

Thus, the author of this thesis tends to disagree with Harvey et al. (2005) when they claim that *"Natural living systems are…dynamical systems designed through natural Darwinian evolution, so it makes sense to consider the artificial equivalent to this design process"*. Should artificial evolution manage to capture all significant elements of Darwinian evolution and achieve complexity similar or equal to it, then this argument might be valid. However, given the simplifications introduced in the way ER is practised, we would choose it as a design methodology over other approaches for many reasons which we have explained previously, but not because it is the artificial equivalent to natural evolution.

Chapter 3

Related work

In this chapter, we provide a literature review of the related work in the areas of the evolution of time-based decision-making mechanisms and communication. Even if the focus of this thesis is on collective behaviours, we also discuss research works considering single agent systems, if the methodology developed there or the results are significant enough and serve as a basis for research performed on multi-agent systems. Thus, section 3.1.3 treats works where one robot is required to make decisions based on its interaction with the environment. Notice that the goal of this section is to provide a high level overview of related work that is common for all experiments to be presented in this thesis; references to the particular literature that is only relevant for a specific experiment will be given inside the chapter in which the experiment is discussed.

3.1 Evolutionary robotics and time-dependent decision-making

In this section we review literature on the issue of evolving time-dependent neural structures. We start with a distinction between non-reactive and reactive control in section 3.1.1, by highlighting the potentialities of the first. Subsequently, in section 3.1.2 we present the neural network type chosen for the experiments detailed in this thesis. Then we present related work in section 3.1.3. Finally, section 3.1.4 treats the issue of downloading time-dependent structures on real robots.

3.1.1 Reactive vs non-reactive control

One way of classifying robotic tasks according to their requirements is to make a distinction between reactive and non-reactive tasks. A reactive task does not require the robots to display any sort of memory of their interactions with the environment or their teammates, while in non-reactive task the opposite holds. Similarly, a non-reactive controller is a controller able to keep memory of previous input patterns while a reactive one can only produce actions based on the current sensory information. A typical example of the latter case is the (multi or single-layer) perceptron, extensively used in ER to tackle tasks where the robots need to produce the same behaviour whenever the same input pattern appears (see Baldassarre et al., 2003; Trianni and Dorigo, 2006, for examples). It is important to stress the fact that a non-reactive controller is a memory-based controller that is at the same time sensitive to the current sensory information.[1]

On the contrary, non-reactive control produces behaviour not only dependent on the current sensory input, but also on past input patterns, either explicitly, or via their influence on internal states. By internal state we mean a state (e.g., the activation state of an internal neuron of the control system of a robot) that can be affected by the previous sensory inputs experienced by the robot and that co-determine, together with the current sensory input, the robot's motor actions. By mediating between perception and actions, internal states can allow agents to produce behaviours that are decoupled from the immediate circumstances while still remaining sensitive to them. Such agents can display complex decision-making that depends on time, and thus they can end up being adaptive than other agents whose behaviour depends only on the current environmental conditions (Beer, 1995).

In this thesis, we use the definition introduced in (Nolfi and Marocco, 2001). Accordingly, a reactive robot is a robot that does not have internal states and for which the current motor action is only dependent on the current sensory state. On the contrary, a non-reactive robot relies both on its internal dynamics and on its current sensory state; motor actions are determined both by sensory information coming from the external environment and internal states.

There exists a variety of neural networks able to display non-reactive behaviour. Some examples are the Elman network (Elman, 1990), the Hopfield network (Hopfield, 1982), spiking neural networks (Maas and Bishop, 1999) and dynamical neural networks (Beer and Gallagher, 1992). In this thesis we exclusively employ dynamical neural networks as controllers for our robots, and in particular the Continuous Time Recurrent Neural Networks (CTRNNs), which is described more in detail in the following section. This is because we believe that dynamical systems theory provides the required mathematical formalisms for the description and the analysis of systems whose behaviour unfolds over time. Thus, time-dependent decision-making in a robot or a robot group can be seen through the perspective of dynamical system theory, because the interactions among robots and between robots and the environment are dynamical interactions.

3.1.2 The CTRNN as a control structure

Continuous Time Recurrent Neural Networks (CTRNNs) have been introduced in evolutionary robotics in (Beer, 1995). They are the reflection of a theoretical approach to cognition which aims to exploit the mathematical tools of dynamical systems theory to investigate issues of interest in adaptive behaviour research. We can say that an agent exhibits adaptive behaviour when it is able to modify its behaviour with respect to changing environmental conditions and/or changes in the behaviour of other agents. According to Beer, there are two fundamental principles which justify the use of the formalism of dynam-

---

[1] Notice that the terms reactive and non-reactive might be misleading in other fields, but in robotics they mean memoryless and memory-based, respectively.

ical systems theory within the context of adaptive behaviour. First, since the fundamental nature of adaptive behaviour in natural systems is to generate the appropriate behaviour at the appropriate time, dynamical systems theory provides the required mathematical formalisms for the description and the analysis of systems whose behaviour unfolds over time. For an embodied agent, time can make the difference between an adaptive behaviour and an unsuccessful one. Second, it looks plausible to consider adaptive behaviour as generated by causal mechanisms which result from the dynamical interactions of elementary units such as cells or molecules, rather than generated by the dynamics of the single elementary units. Thus, in an effort to explain adaptive behaviour, we need to resort to the structure of the internal dynamics; dynamical systems theory is the right tool to give us insight into such processes.

CTRNNs represent a convenient way of instantiating a dynamical system to control the behaviour of autonomous robots. CTRNNs differ from the classical artificial neural networks (e.g., the perceptron), as they can be expressed as a system of differential equations. Each node within a CTRNN has its own state—the activation level—whose rate of change is specified by a time constant associated with each node. Furthermore, the nodes within the network are self-connected, as well as interconnected in an arbitrary way with each other. These two features allow the network to develop dynamical behaviour in which the state of nodes alters the behavioural output of the system even if the sensory input remains constant.

CTRNNs are arguably the simplest nonlinear, continuous dynamical neural network model (Beer, 1995)[2]; however, despite their simplicity, they are universal dynamics approximators in the sense that, for any finite interval of time, CTRNNs can approximate any smooth dynamical system arbitrarily well (see Funahashi and Nakamura, 1993, for details). Finally, fixed weight CTRNNs have been demonstrated to be able to produce learning behaviour in robots (see Tuci et al., 2002); this goes against the belief that learning in ANNs necessarily corresponds to weight changes (plastic weights).

3.1.3 The evolution of time-dependent structures

Several studies have described evolutionary simulation models in which time-dependent structures are evolved to control the behaviour of robots required to make decisions based on their experiences.

The evolution of time-dependent structures and decision-making mechanisms has been extensively studied on the T-maze problem (see Ziemke and Thieme, 2002; Blynel and Floreano, 2003). The robot is required to find its way to a goal location, placed at the bottom of any of the two arms of the maze. When at the T-junction, the robot has to decide whether to turn left or right. The correct decision can be made if the agent is capable of either exploiting perceptual cues which were available to it while it was navigating down the first corridor, or capable of retaining in its memory something about previous trials in a similar T-maze. In (Ziemke and Thieme, 2002), weight changing

---

[2] Notice that this system is in fact a discrete time network (the network is integrated using the forward Euler integration method); however, the term Continuous Time Recurrent Neural Network is commonly used in the literature and will also be used in this thesis.

mechanisms provide the agents the required plasticity to exploit the relationship between the location of light signals placed roughly in the middle of the first corridor, and the turn to make at the junction. Blynel and Floreano (2003) allow the agent to experience the environment in a first trial, in which the success or failure play the role of a reinforcement signal, in order to associate the position of the goal with respect to the T-junction. The work illustrated in (Nolfi, 2002) investigates a discrimination task in which a robot, while navigating through a maze, must recognise if it is located in one room or another. Here, the agent exploits environmental cues, such as navigating through subsequent corners of the maze, and fine-tuned time-dependent structures to take the correct decision.

In (Tuci et al., 2002), evolved CTRNNs provide the agents the required plasticity to discover the spatial relationship between the position of a landmark and the position of a goal. In this study, the spatial relationship between the goal and the landmark can be learnt by "remembering" from previous trials the relative position of the landmark with respect to the goal. Notice that in this work, which is revisiting the task introduced in (Yamauchi and Beer, 1994), individual learning is produced by a network with fixed weights and without weight changing mechanisms. Also, contrary to (Yamauchi and Beer, 1994), a single integrated (i.e., non-modularised) neural network managed to produce both reactive and non-reactive behaviour.

In the studies reviewed above, the discrimination is based on the recognition of distinctive environmental contingencies and the maintenance of these experiences through time, as a form of short term memory. However, there are works in the literature in which the cue which allows the agent to make the discrimination has to deal with the persistence over time of a perceptual state common to all elements to be distinguished, rather than with the nature of the cue itself employed to make the discrimination. That is, due to the nature of the agent's sensory apparatus, the different types of environment can be distinguished solely because a perceptual state might be perceived by the agent for a longer time in one environment than in the others.

Nolfi and Marocco (2001) define agents that exploit internal representations as well as information directly available from their sensors and that are able to extract their internal representations autonomously by interacting with the environment, as agents that are able to integrate sensory-motor information over time. These agents rely on a mixed strategy in which basic sensory-motor mechanisms are complemented and enhanced with additional internal mechanisms (see Nolfi and Marocco, 2001). The authors solved a navigation and a visual discrimination task in changing environmental conditions with non-reactive controllers—the architecture was feed-forward, but there were hidden neurons with recurrent connections, and thus this network was able to take time into account—and proved that these controllers outperform reactive controllers when faced with the same task. Moreover, they showed that *"more complex individuals that are able to integrate information over time still rely on the same sensory-motor strategy adopted by reactive individuals and do not discover a completely different strategy."*

Beer (2003b) applied the perspective and tools of dynamical systems theory on the study of active categorical perception. In particular, he evolved a CTRNN controller for a simulated agent that had to catch circular objects and to avoid diamond-shaped ones. Successful agents determined the object width in order to perform the visual discrimination and the decision to catch or avoid was a process spanning over time rather than being a discrete event. The implications of this study on cognitive science were very significant, since the article adopted a significant anti-representational point of view.

> After close examination of the circle-catching, diamond-avoiding agent, we find no circle or diamond detectors and nothing that resembles a representation of a circle or a diamond. Appeals to internal representations are not needed to explain the agent's behaviour, whereas the right set of dynamical equations allows us to understand what is going on... It follows then that representations cannot explain perception and action, because, under Beer's view, these are properties of the coupled agent/environment system and not properties of the internal machinery of the agent.
>
> (De Pinedo and Noble, 2003)

Tuci et al. (2004) designed decision-making mechanisms for an autonomous robot equipped with simple sensors, which integrates over time its perceptual experience in order to initiate a simple signalling response. Contrary to other previous similar studies, in this work the decision-making was uniquely controlled by the time-dependent structures of the agent's controller (a fully connected CTRNN), which in turn, are tightly linked to the mechanisms for sensory-motor coordination. The results of this work showed that a single dynamical neural network, shaped by evolution, makes an autonomous agent capable of "feeling" time through the flow of sensations determined by its actions. Further analysis of the evolved solutions revealed the nature of the selective pressures which facilitate the evolution of fully discriminating and signalling agents. This research work served as the basis of our experimental setup described in chapter 4. In section 4.2 we will highlight similarities and differences between the two experiments.

Finally, a recent research work detailed in (Gigliotta and Nolfi, 2008), describes the results of an experiment where an e-puck[3] robot solves the double T-maze problem and can memorise the position of the goal. By integrating sensory information over time, evolved robots generate different internal states associated with the part of the environment where the robot is located. These states are similar when the robot visits or re-visits the same environmental location.

3.1.4 Porting non-reactive controllers to real robots

Due to the number of trials needed to test individuals, the design of robot controllers by means of artificial evolution is usually carried out by using simulation models. However, the digital medium might fail to take into account phenomena that are important for the functional properties of the evolved controllers. As a consequence, controllers evolved in simulation might be less effective in managing real-world sensing and actuation (see also Matarić and Cliff, 1996). One of the main contributions of our work is to show that evolved CTRNNs successfully control real robots. This is a practice that has to be taken

---

[3] See `http://www.e-puck.org/`.

into account to assure that the behaviours we want our robots to display are viable and observable in the real world and not only in a simulated environment.

There exist several works in the literature that deal with porting an Artificial Neural Network (ANN) able to display memory to reality. Paine and Tani (2005), Blynel and Floreano (2003), and Jakobi (1997) all port evolved CTRNNs onto real Khepera robots, but although the networks used are non-reactive, the tasks described—variations of the T-maze—are in essence solved by switching through reactive strategies (see Ziemke and Thieme, 2002). Urzelai and Floreano (2001) downloaded a PNN (Plastic Neural Network) on a real Khepera, but the solution to the task is also reactive. Quinn et al. (2002) report on work done on real hardware—on a collective task, but the network they use is based on model spiking neurons. Recently, Gigliotta and Nolfi (2008) successfully downloaded a network able to display internal dynamics (it included internal neurons implemented as leaky integrator neurons) to a real e-puck robot. The task involved a double T-maze and a robot that had to memorise the location of the target by employing integration over time of sensory input.

To the best of our knowledge, there is no work in the literature treating the issue of porting a CTRNN to a real robot for a task that requires the integration over time of the robot's perception. Moreover, apart from the work of Quinn et al. (2002), there is no work involving the porting of non-reactive controllers on a group of real robots (for a collective task).

3.2 Evolutionary robotics and collective robotics - The evolution of signalling and communication

In this section we review research work focused on the issue of designing through artificial evolution neural mechanisms for communicative behaviour in groups of autonomous robots. Consequently, we do not consider those interesting works on communication in multi-robot systems in which the mechanisms for social interactions are designed by using other methods than ER. For a survey of work in those fields, we point the reader to the following articles: Balch and Arkin (1994); Cao et al. (1997); Fong et al. (2002); Støy (2001). We also do not mention here works that treat the evolution of communication by considering conflicts of interest; the focus of the section is to review related work studying the evolution of cooperative behaviour in autonomous agents.

The first works in the evolution of communication in simple machines, animats or robots date back to the beginning of the 90's. MacLennan and Burghardt (1993) report on a series of experiments whose aim was to evolve cooperative communication in a population of simple machines. The non-embodied agents (called simorgs), are limited to their local environment, while they are able to modify a global environment that can be sensed by all simorgs. Selection acts on cooperative simorgs and thus there is a priori introduced selective pressure for cooperative communication. The control structure of the simorgs is a finite-state machine (rule table). The results show that if communication is allowed, the average fitness of the population increases much faster than if communication is suppressed. Moreover, should the use of communication be enhanced by using a simple learning module, the average fitness increase is even faster. This work experimentally proves the relation between learning and communication; however, the agents considered are not embodied, and the local and global environments are separated. Also, learning is not co-evolving with communication but simply considered on/off.

Another important work is the research work of Werner and Dyer (1992). This is where the idea that there should be no direct pressure on communication was introduced. More specifically, the authors argue that the fitness of individuals should not depend on the communication protocol eventually used, but on how well they solve a “natural” task. This way, evolution is not constrained by the fitness function and it is free to shape “creative” solutions that rely on unbiased fitness function design (see also chapter 2, section 2.7.2.1 and the discussion on autonomous decision-making). Evolved neural networks are used as control structures of male and female agents. The task was to co-evolve a population of embodied simulated males and females who live in a grid world and can agree on the interpretation of signals emitted only by the females. Females are supposed to guide the behaviour of perceived males, while the latter are blind and can only perceive the signals emitted by the females. Indeed, communication evolves and the authors also notice the evolution of different, competing dialects. In contrast to the work of MacLennan and Burghardt (1993), this research work is considering a rather realistic task, where signals have a meaning in the context of the task considered and are not simply abstract associations.

Di Paolo (2000) performed research on the evolution of acoustic communication between two agents that have to use sound signals in order to localise and stay close to each other, in the absence of other sensory proximity information. Employing evolved dynamical neural networks as control units, coordination is achieved through one sound channel via coupled interactions that resemble turn-taking.

In the seminal work on the evolution of communication, Quinn (2001); Quinn et al. (2003) managed to achieve the evolution of communication without dedicated channels. The author(s) evolved neural networks controlling a team of mobile robots for the ability to move by remaining close to one another. Evolved individuals are able to solve the coordination problem by communicating through a sequence of sensory-motor interactions. The evolved neuro-controllers were also tested on real robots (see Quinn et al., 2002).

Still in the domain of social interactions for collective navigation tasks, Baldassarre et al. (2003) evolved reactive neuro-controllers for a group of homogeneous simulated robots required to move together towards a target. Contrary to the work described in (Quinn et al., 2003), Baldassarre et al. made use of a dedicated communication channel in the form of a loudspeaker continuously emitting a tone and directional microphones. Similar communication specifications that are again hardly portable on real robotic hardware (at least on the *s-bot* which was used in these works) were used in (Dorigo et al., 2004), where a reactive controller is evolved to control simulated robots whose goal is to aggregate and move straight together.

In (Marocco et al., 2003), the authors describe a research work performed in a simulated environment, where a robotic arm has to categorise the environment; in detail, the arm has to discriminate between objects with different shapes by exploring them with its sensory apparatus. Communication is not represented in the fitness function and its emergence is open to evolution. Indeed, it emerges if it concerns parents and offspring. That is, communication emerges when it is the parents of the individuals that provide information on the object that their offspring has to actively explore and categorise. It is also shown that an incremental approach—first evolving the individual categorisation mechanisms and then allowing for communication—facilitates the emergence of communication. An important conclusion is that already evolved cognitive structures (as categorisation) can serve as prerequisites of communication. However, notice that communicating agents are not sharing the same environment—they are parents and offspring.

Marocco and Nolfi (2006a) report on the results of a collective navigation task, where the cooperation among the agents is required in order to accomplish their objective. However, there is no explicit selective pressure on communication and no biases on potential uses of acoustic signals that the agents can use to communicate. It is important to notice that the authors assume (i) that the robots can distinguish between *self* and *non-self* sound components, that is, sound emitted by themselves and sound emitted by other robots and (ii) that the robots are able to distinguish four different directions of sound. These choices make the implementation of evolved mechanisms on real hardware rather impractical.

Wischmann and Pasemann (2006) present an experiment where robots sharing the same environment are looking for food patches. Even if communication is not explicitly rewarded, it emerges since it enhances the group's behaviour. The authors utilise an incremental approach in which communication emerged only for a very simple environment but its use is conserved as the environmental complexity increases. Controllers are evolved in a realistic simulation but were not transferred on real hardware.

Trianni and Dorigo (2006) managed to evolve signalling behaviours in a collective hole avoidance task. Once again, without explicit selective pressure on communication, if robots controlled by simple perceptrons are provided with signalling capabilities, acoustic communication can emerge since it enhances the performance of the group that otherwise has to rely solely on physical interactions to solve the reactive task at hand. The evolved communication protocol was tested on real hardware.

A very important result of this last work is that a fully evolved approach leads to better results than an approach where the acoustic communication system is hard-coded into the robots' controllers. This shows that the automatic process can detect features of the robot-environment interaction space that may be adaptive or beneficial for the robots; these features may not always be a priori evident to the experimenter. As we stressed in chapter 2, section 2.7.2.3, it is difficult or even impossible for an external observer to predict the dynamic interactions among robots and between robots and environment.

Chapter 4

Experiment I: categorisation and communication

In this chapter, we report on a series of experiments performed both in simulation and on real hardware, concerning the relationship between autonomous decision-making and the evolution of communication, in the context of a categorisation task. These experiments bring the problem of decision-making together with the interest in self-organising communicating systems to a real-world scenario, in which the switch from individual to collective behaviour via an emergent communication protocol can be empirically investigated. In particular, this switch is governed by time-based decision-making structures that integrate over time sensory information available to the robots. Moreover, we look at issues directly implicated in the switch from solitary to collective behaviour, such as the emergence of a communication system and its relation to the individual decision-making.

In our experimental setup, the shaping of individual mechanisms for categorisation will be the result of the evolutionary process; we will not impose the way the robot should perform the categorisation (as for example in providing a Hebbian learning rule). Furthermore, even if the robots will be provided with signalling capabilities, that is, a sensor to perceive sound and an actuator to emit sound, the use of those signalling capabilities will be entirely left to evolution. That is, potential signalling protocols are not pre-designed by the experimenter. On top of those two levels of autonomous decision-making, there is another level, which concerns the choice between solitary and social behaviour; our experimental setup will not impose the use of communication to the system.

In section 4.1 we lay down the motivations and the biological background for our study and in section 4.2 we relate our work to research work in the literature on time-based decision-making and to research works treating the evolution of communication in embodied agents. The experimental setup and the description of the task will follow in section 4.3; results in simulation are presented in section 4.4, while results of experiments performed on real hardware are presented in section 4.5; the chapter closes with discussion and conclusions (sections 4.7 and 4.8).

4.1 Biological background and motivations

Several research works in zoology, and in particular in social foraging, have reported that the foraging behaviour of animals changes if the animals are situated in a social context. For example, Elgar (1987) shows that social companionship in house sparrows leads to higher feeding rates, as each individual eventually spends less time scanning for predators. Similarly, Fernandez-Juricic et al. (2005) show that while foraging, starlings spend more time scanning for predators once social information is reduced. A general problem common to biology and robotics concerns the definition of the mechanisms necessary to decide (i) when it is better to pursue a particular action in a certain location and when it is better to leave for pursuing a similar or a different activity in a similar or different location; and (ii) whether it is better to pursue solitary actions or to initiate cooperative strategies. This problem is not limited to foraging alone, but it extends to many activities a natural or artificial agent is required to carry out. Autonomous agents may be asked to change their behaviour in response to the information gained through repeated interactions with their environment. For example, in a group of robots, although many individual actions might be simpler to carry out than a single coordinated activity, they might result less efficient (see Trianni et al., 2004).

In our study, we focus on the evolution of communication in the form of a simple signalling system. Nature abounds with examples from social species, where simple (compared to human communication and language) signalling mechanisms are used. For example, the alarm calls of vervet monkeys given with respect to the type of predator approaching have been studied in depth (see Struhsaker, 1967; Seyfarth et al., 1980, for examples). Alarm calls are also observed in bird species, squirrels, etc. (Hauser, 1997; Sherman, 1977). Food calls are another example of cooperative signalling. Animals like chimpanzees attract conspecifics once they discover food resources. The dance of the honey bee is possibly the most elaborate and striking example (Von Frisch, 1967).

Since Darwin, scientists have been trying to explain the evolution of such altruistic signals in animal societies. Ethologists justified the existence of such cooperative and honest signalling by invoking group selection theory: animals behave in such ways so to maximise the benefit of the group or the species (see Tinbergen, 1964, for example). However, the alternative of kin selection was presented (Hamilton, 1964) and the naïve application of group selection as an explanation was shown to be unwarranted (Williams, 1966; Dawkins, 1976). Kin selection suggests that animals can behave with apparent altruism towards conspecifics since this can be to their own long-term genetic benefit.

Game theoretical models in the 1970s and 80s mathematically demonstrated that cheating strategies will normally invade populations of honest signallers (Maynard-Smith, 1982). Thus the interest of researchers focused on how to identify conditions that can lead to the emergence of stable cooperation, as, for example, Hamilton's kin-selection theory (Hamilton, 1964), reputation-based models (Nowak and Sigmund, 1998; Nowak, 2006), or the effect of topology (Santos et al., 2006).[1] The game theoretical models studying such issues typically consider signalling capabilities that are built into the agent's behavioural

---

[1] There is a very large amount of literature on these issues, but since it falls out of the scope of this thesis, we will not refer to any more works.

capacity. Thus, they do not allow the investigation of the origin of signalling behaviour; their focus is on the study of the conditions for the stability of communication rather than on its origin (Noble, 1999).

The experimental setup we use in this work differs in several aspects from these game theoretical models. First, we are attempting to study the origin of signalling, since signalling capabilities are evolved (sensors and actuators are available for communication, but there is no requirement that the robots use them). More specifically, we discuss the existence of possible cues that served as precursors for the signals employed by our robots, through the process of ritualisation. Second, in our work the possibility of cheating and dishonest signalling is excluded, because the evaluation of the fitness of a group of individuals is done at the group level and the individuals composing the group are genetically identical clones. Our aim is to understand how communication may emerge in a robotic system, in the absence of explicit selective pressures. In other words, we aim to understand the conditions under which a group of agents will switch to social behaviour, and the implications of that switch for the performance of the group in a certain scenario. Our focus is more on the evolution of signalling than the evolution of cooperation. Our implementation has been influenced by an ethological perspective, even though this does not mean that we are trying to do robot ethology.

4.2 Related work

The goal of this section is to highlight similarities and differences between our approach and our experimental setup and related work in the areas of evolving time-based decision-making, porting such structures on real robots and the evolution of signalling behaviours in autonomous robots. This will be done in an effort to bring together works which treat time-dependent decision-making and social behaviour.

Our approach shares some characteristics with some of the works presented in chapter 3, section 3.2. More specifically, we agree with the ideas developed in (Werner and Dyer, 1992), concerning the absence of the representation of communication in the fitness function. Like that, the automatic process can shape the communication in an unbiased way. Moreover, as in (Wischmann and Pasemann, 2006), communication is not strictly required to solve the task at hand. The choice between social and solitary behaviour is left to evolution. However, while in (Wischmann and Pasemann, 2006) social behaviour is expected to lead to higher fitness values, in our work this is a priori not clear—should there be any benefit for social behaviour, this should be "discovered" by the evolutionary process. Also, we want communication and individual cognitive capabilities to co-evolve, as in (Marocco and Nolfi, 2006a). This might help to shed light on the origins of signalling behaviours using a rather prejudice-free approach, while in incremental approaches (see Wischmann and Pasemann, 2006; Marocco et al., 2003, for example) the a priori introduced sequential nature of communication and, for example, individual categorisation can significantly bias the evolved communication protocol. Thus, our approach can be called integrated and non-incremental.

Finally, our work will focus on solving a real-world task, that is, we will put particular emphasis on porting the evolved signalling strategies on real hardware. This is a focus that might constrain the space of potential evolved protocols. For example, in (Marocco and Nolfi, 2006a; Wischmann and Pasemann, 2006) complex signalling protocols emerge, but their feasibility on real hardware is questionable. The same goes for the experiments performed in simulation and presented in (Baldassarre et al., 2003; Dorigo et al., 2004); the evolution of a simple signalling mechanism for robot aggregation and formation movement is documented, but the use of directional microphones makes the communication specifications hardly portable on real robotic hardware (at least on the *s-bot* which was used in these works). On the contrary, the signalling mechanisms used in (Trianni and Dorigo, 2006) are extremely simple (beeps) and portable on real robots. Our approach will make use of the same signalling capabilities used in the latter research work, thus a simple beep that is emitted by any robot and that is perceived by the rest of the team. Mechanisms for distinguishing between *self* and *non-self* components will not be provided to the robots and their evolution is not the goal of this research work.

Concerning the porting of the evolved neuro-controllers on real robotics hardware, we find it important to repeat what was already said in chapter 3, section 3.1.4. To the best of our knowledge, there is no work in the literature treating the issue of porting a CTRNN to a real robot, for a task that requires the integration over time of the robot's perception. In this respect, it is worth noting that the decision-making mechanism relies on the continuum of the sensory information (i.e., how the sensory inputs unfold in time) in order to determine subsequent actions. Our experimental setup described in chapter 4.3 treats a task that requires the integration of sensory information over time for a task very different than the T-maze, where the robots' decisions will unfold in time. Moreover, the task will be collective, that is, there is more than one robot involved, thus interactions among robots of the group will take place. Apart from the work of Quinn et al. (2002), there is no work involving the porting of non-reactive controllers on a group of real robots.

Overall, the main challenges in porting our controllers to reality are: (i) the possible disruptive effects on the evolved time-dependent mechanisms caused by the sensor/actuator noise present in reality, and (ii) the potential inter-robot differences (e.g., subtle hardware differences, different noise on sensors/actuators, etc.) that might cause robots to behave differently and that might have negative effects on the individual performance or even on the performance of the whole group.

4.3 Methods

This section is structured as follows: first we describe the task the robots should solve in section 4.3.1, then we go on to discuss the simulation model used (see section 4.3.2) and the controller and the evolutionary algorithm chosen (see section 4.3.3), and we finish by describing the fitness function in section 4.3.4.

4.3.1 Description of the task

The task we consider is a categorisation task in which two robots are required to discriminate between two different environments using temporal cues, that is, by integrating their

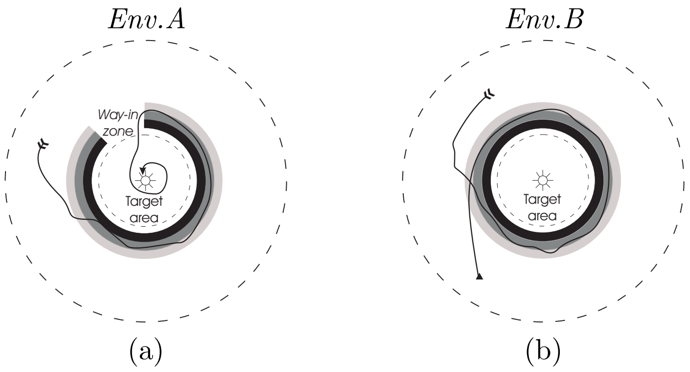


Figure 4.1: The task. (a) *Env.A* is characterised by the *way in* zone. The *target area* is indicated by the dashed circle. (b) In *Env.B* the *target area* cannot be reached. The continuous arrows are an example of a good navigation strategy for one robot.

perceptual inputs over time. At the start of each trial, two simulated robots are placed in a circular arena with a radius of 120 cm (see figure 4.1), at the centre of which a light bulb is always turned on. The robots are positioned randomly at a distance between 75 and 95 cm from the light, with a random orientation between −120° and +120° with respect to the light. The robots perceive the light through their ambient light sensors. The colour of the arena floor is white except for a circular band, centred around the lamp covering an area between 40 and 60 cm from it. The band is divided in three sub-zones of equal width but coloured differently: light gray, dark gray, and black. Each robot perceives the colour of the floor through its floor sensors, positioned under its chassis. Robots are not allowed to cross the black edge of the band close to the light. This black edge can be seen as a circular trough that prevents the robots from reaching the light. The coloured zones can be seen as an indication of how close the robots are to the "danger". There are two types of environment. In one type—referred to as *Env.A*—the band has a gap, called the *way in* zone, where the floor is white (see figure 4.1a). In the other type, referred to as *Env.B*, the band completely surrounds the light (see figure 4.1b). The *way in* zone represents the path along which the robots can safely reach the *target area* in *Env.A*—an area of 25 cm around the light. In contrast, the robots cannot reach the proximity of the light in *Env.B*, and in this situation their goal is to leave the band and reach a certain distance from the light source. Robots have to explore the arena, in order to get as close as possible to the light. If they encounter the circular band they have to start looking for the *way in* zone in order to continue approaching the light, and once they find it, they should get closer to the light and remain in its proximity for 30 sec. After this time interval, the trial is successfully terminated. If there is no *way in* zone (i.e., the current environment is an *Env.B*), the robots should be capable of "recognising" the absence of the *way in* zone and leave the band by performing antiphototaxis.

Each robot is required to use a temporal cue in order to discriminate between *Env.A* and *Env.B*, as in (Tuci et al., 2004). This discrimination is based on the persistence of the perception of a particular sensorial state (the floor, the light or both) for the amount of

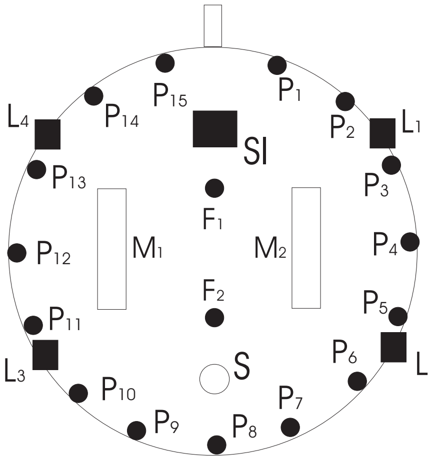


Figure 4.2: Sensors and motors of the simulated robot. The robot is equipped with four ambient light sensors (L_1 to L_4), two floor sensors F_1 and F_2, 15 proximity sensors (P_1 to P_{15}) and a binary sound sensor, called SI (see text for details). The wheel motors are indicated by M_1 and M_2. S is the sound signalling system (loudspeaker).

time that, given the trajectory and speed of the robot, corresponds to the time required to make a loop around the light. The integration over time of the robots' sensorial inputs is used to trigger antiphototaxis in *Env.B*. Communication is not required to solve the task considered. In particular, the fitness function we use does not explicitly reward the use of signalling, in contrast with (Tuci et al., 2004).[2] However, robots are provided with a sound signalling system that can be used for communication. The emergence of a signalling convention by which the robots can affect each other's behaviour is entirely open to the dynamics of the evolutionary process. This issue is further discussed in section 4.4.

4.3.2 The simulation model

The controllers are evolved in a simulation environment which models some of the hardware characteristics of the *s-bots* (see section 2.3.1). In this work, we make use of four ambient light sensors, placed at $-112.5°$ (L_1), $-67.5°$ (L_2), $67.5°$ (L_3), and $112.5°$ (L_4) with respect to the *s-bot*'s heading, fifteen infra-red proximity sensors placed around the turret (P_1 to P_{15}), two floor sensors F_1 and F_2 positioned facing down on the underside of the robot with a distance of 4.5 cm between them, and an omni-directional sound sensor SI (see figure 4.2). The motion of the robot implemented by the two wheel actuators (M_1 and M_2) is simulated by the differential drive kinematics equations, as presented in (Dudek and Jenkin, 2000), and a loudspeaker S is available for possible signalling. Light and proximity sensor values are simulated through a sampling technique (Miglino et al., 1995). The robot floor sensors assume the following values: 0 if the sensor is positioned over white

---

[2] Notice that in the task described in (Tuci et al., 2004) there was only one robot in the environment; the signal emitted by the robot was used as a means to denote to the experimenter that the robot has decided it is situated in *Env.B*.

floor; $\frac{1}{3}$ if the sensor is positioned over light gray floor; $\frac{2}{3}$ if the sensor is positioned over dark gray floor; 1 if the sensor is positioned over black floor. The loudspeaker produces a binary output (on/off); the sound sensor has no directionality or intensity features. During evolution, 10% random noise was added to the light and proximity sensor readings, the motor outputs and the position of the robot. We also added noise of 5% to the reading of the two floor sensors, by randomly flipping between the four aforementioned values. No noise was added to the sound sensor.[3]

4.3.3 The controller and the evolutionary algorithm

We use fully connected, thirteen neuron CTRNNs (Beer and Gallagher, 1992, see figure 4.3 for a depiction of the network and chapter 3, section 3.1.2 for more details). All neurons are governed by the following state equation:

$$\frac{dy_i}{dt} = \frac{1}{\tau_i}\left(-y_i + \sum_{j=1}^{13} \omega_{ji}\sigma(y_j + \beta_j) + gI_i\right), \qquad \sigma(x) = \frac{1}{1+e^{-x}} \tag{4.1}$$

where, using terms derived from an analogy with real neurons, τ_i is the decay constant, y_i represents the cell potential, ω_{ji} the strength of the synaptic connection from neuron j to neuron i, $\sigma(y_j + \beta_j)$ the firing rate, β_j the bias term, g the gain and I_i the intensity of the sensory perturbation on sensory neuron i. The connections of all neurons to sensors and actuators is shown in figure 4.3. Neurons N_1 to N_8 receive as input a real value in the range [0,1]. Neuron N_1 takes as input $\frac{L_1+L_2}{2}$, $N_2 \leftarrow \frac{L_3+L_4}{2}$, $N_3 \leftarrow F_1$, $N_4 \leftarrow F_2$, $N_5 \leftarrow \frac{P_1+P_2+P_3+P_4}{4}$, $N_6 \leftarrow \frac{P_5+P_6+P_7+P_8}{4}$, $N_7 \leftarrow \frac{P_9+P_{10}+P_{11}+P_{12}}{4}$ and $N_8 \leftarrow \frac{P_{13}+P_{14}+P_{15}}{3}$. Neuron N_9 receives a binary input (i.e., 1 if a tone is emitted by either agent, 0 otherwise) from the microphone SI, while neurons N_{10} to N_{13} do not receive input from any sensor. The cell potentials (y_i) of N_{11} and N_{12}, mapped into [0,1] by a sigmoid function (σ) and then linearly scaled into [-4.0 $\frac{cm}{s}$, 4.0 $\frac{cm}{s}$], set the robot motors output. It is important to mention that the speed that these values translate to is not the maximum possible speed of the robot, but only half of it. This is due to the fact that after some initial experimentation, we found that if we use a faster robot, we have a higher chance of getting a false reading from the floor sensors and in general a worse sensory-motor coordination. The cell potential of N_{13}, mapped into [0,1] by a sigmoid function (σ) is used by the robot to control the sound signalling system (the robot emits a sound if $y_{13} \geq 0.5$). The parameters ω_{ji}, τ_i, β_j and g are genetically encoded. Cell potentials are set to 0 when the network is initialised or reset, and circuits are integrated using the forward Euler method with an integration step-size of 0.1.

A simple generational genetic algorithm (GA) is employed to set the parameters of the networks (Goldberg, 1989). The population contains 100 genotypes. Each genotype is a vector comprising 196 real values (169 connections, 13 decay constants, 13 bias terms, and

---

[3] The reason for this last choice is the fact that the sound sensor proved to be 100% reliable in reality. Of course, adding noise to the sound sensor would force the simulation to address the issue of the reliability of the evolved signals and thus produce neural mechanisms able to cope with noisy communication. This issue, while an interesting one, is beyond the scope of our research work.

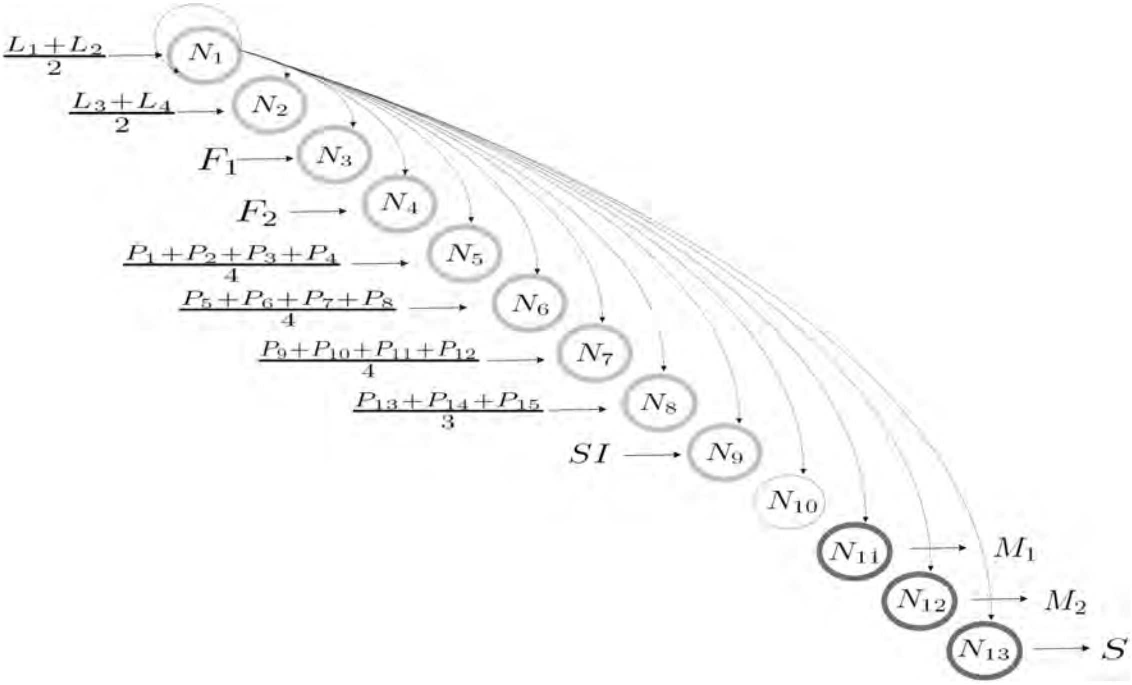


Figure 4.3: The fully connected CTRNN architecture. Neurons are represented as circles. Circles with the light gray outline represent the input neurons, while circles with the heavy gray outline represent the output neurons. Only the efferent connections for N_1 are drawn: all other neurons are connected in the same way. We show for all input neurons the combination of sensors that serve as inputs, and for all output neurons the corresponding actuator. N_{10} is not connected to any sensor or actuator.

a gain factor). Initially, a random population of vectors is generated by initialising each component of each genotype to values chosen uniformly random in the range [0,1]. Subsequent generations are produced by a combination of selection with elitism, recombination and mutation. For each new generation, the three highest scoring individuals ("the elite") from the previous generation are retained unchanged. The remainder of the new population is generated by fitness-proportional selection from the 70 best individuals of the old population. New genotypes, except "the elite", are produced by applying recombination with a probability of 0.1 and mutation. Mutation entails that a random Gaussian offset is applied to each real-valued vector component encoded in the genotype, with a probability of 0.15. The mean of the Gaussian is 0, and its standard deviation is 0.1. During evolution, all vector component values are constrained within the range [0,1]. Genotype parameters are linearly mapped to produce CTRNN parameters with the following ranges: biases $\beta_j \in$ [-2,2], weights $\omega_{ji} \in [-6, 6]$ and gain factor $g \in$ [1,12]. Decay constants are firstly linearly mapped onto the range $[-0.7, 1.7]$ and then exponentially mapped into $\tau_i \in [10^{-0.7}, 10^{1.7}]$. The lower bound of τ_i corresponds to a value slightly smaller than the integration step-size used to update the controller; the upper bound corresponds to a value slightly bigger than the average time required for a robot to reach and perform a complete loop around the band in shades of gray.

4.3.4 The fitness function

During evolution, each genotype is coded into a robot controller, and is evaluated for 10 trials, 5 in each environment. Both robots in the ten trials have the same controller, that is, we use a homogeneous system. The sequence order of environments within the ten trials does not influence the overall performance of the group since each robot controller is reset at the beginning of each trial. Each trial differs from the others in the initialisation of the random number generator, which influences the robots' starting positions and orientation, the position and amplitude of the *way in* zone (between 45° to 81°), and the noise added to motors and sensors. Within a trial, the robot life-span is 100 sec (1,000 simulation cycles). The final fitness attributed to each genotype is the average fitness score of the 10 trials. In each trial, the fitness function E is given by the following formula:

$$E = \frac{E_1 + E_2}{2 \times (n_c + 1)},$$

where n_c is the number of (virtual) collisions in a trial, that is the number of times the robots get closer than 2.5 cm to each other (if $n_c > 3$, the trial is terminated) and E_i, $i = 1, 2$, is the fitness score of robot i, calculated as follows:

- If the trial is in *Env.A*, or the robot in either environment has not yet touched the band in shades of gray or crossed the black edge of the band, then its fitness score is given by $E_i = \frac{d_i - d_f}{d_i}$.
- Otherwise, that is if the band is reached in *Env.B*, $E_i = 1 + \frac{d_f - 40}{d_{max} - 40}$.

d_i is the initial distance of the robot to the light, d_f is the distance of the robot to the light at the end of the trial and $d_{max} = 120$ cm is the maximum possible distance of a robot from the light. In cases where $robot_i$ ends up in the *target area* in *Env.A*, we set $E_i = 2$. From the above equations we can see that this is also the maximum value of E_i that a robot can obtain in *Env.B*, which corresponds to the robot ending up 120 cm from the light ($d_f = 120$). So if both robots are successful, the trial gets the maximum score of 2.

Notice that the design of the fitness function does not interfere with the individual strategies that successful robots should employ in order to solve the task at hand. The robots are simply demanded to perform the discrimination between the two environments. Moreover, the choice between social and solitary behaviour is also left entirely to evolution, since this fitness function rewards agents that develop successful discrimination strategies and end up doing the correct action in each environment, regardless of any use of sound signalling. That is, a genotype that controls a group that solves the task without any signalling or communication can in principle achieve the same fitness score as one that makes use of communication. Should evolution come up with a communication protocol, this protocol is not predefined by the experimenter. The only level where our experimental setup constrains the way evolution can make use of the sound signalling system is of course the level of the hardware, which cannot evolve.

4.4 Results: simulated agents

In this section, we introduce the evolutionary simulations that we ran in order to obtain controllers that solve the task introduced in the previous sections and we then present a series of post-evaluation tests concerning simulated robots, aimed to unveil operational principles of the evolved behaviour. In section 4.4.1, we show that sound signalling is a functional element of the behavioural strategies in the majority of successful groups of robots.

Twenty evolutionary simulation runs, each using a different random initialisation, were run for 12,000 generations. Thirteen evolutionary runs produced successful groups of robots. Note that a group is successful if both robots approach the band and subsequently (i) reach the *target area* through the *way in* zone in *Env.A*; (ii) leave the band performing antiphototaxis in *Env.B*. We arbitrarily demand that the successful accomplishment of this task corresponds to an average fitness score $F \geq 1.8$. In those seven evolutionary runs considered not successful, the fitness score recorded during the evolutionary phase by the best groups at each generation was always lower than 1.8. For each successful run, we chose to post-evaluate the best group of each generation whose fitness score was higher than 1.8.

The post-evaluation tests are meant to provide a better estimate of the behavioural capabilities of these groups. In fact, the fitness of the best evolved controllers during evolution might have been an overestimation of their ability to guide the robots in the task. In general, the best fitness scores take advantage of favourable conditions, which are determined by the existence of inter-generational variation in starting position and orientation and other simulation parameters. The entire set of post-evaluations should establish whether the groups chosen from the thirteen successful runs can effectively solve the task and at the same time ascertain whether signalling behaviour characterised the successful strategies. We employed the average fitness score F over a set of 500 trials in each type of environment as a quantitative measure of the effectiveness of the evolved groups' strategy.

Table 4.1 shows, for each successful evolutionary run (i), the results of the best group among those chosen for post-evaluation. These groups are referred to as g_i. We can notice that all these groups achieve an average fitness score in each environment higher than 1.8 (see table 4.1 columns 2, 3, 6, and 7). Thus, they proved to be particularly successful in performing the task. The post-evaluation tests also reveal that among the successful groups, nine groups (g_1, g_2, g_5, g_6, g_7, g_8, g_9, g_{13}, g_{19}) make use of sound signalling. In particular, the use of sound strongly characterises the behavioural strategies of the groups when they are located in *Env.B*. In *Env.A* signalling is, for all these groups, rather negligible—see table 4.1 columns 4, 5, 8, and 9, which refer to the average percentage and standard deviation of the time either robot emits a signal during a trial. In groups g_{10}, g_{14}, g_{16}, g_{18}, the robots do not emit sound during post-evaluation in either environment.

4.4.1 Sound signalling and communication

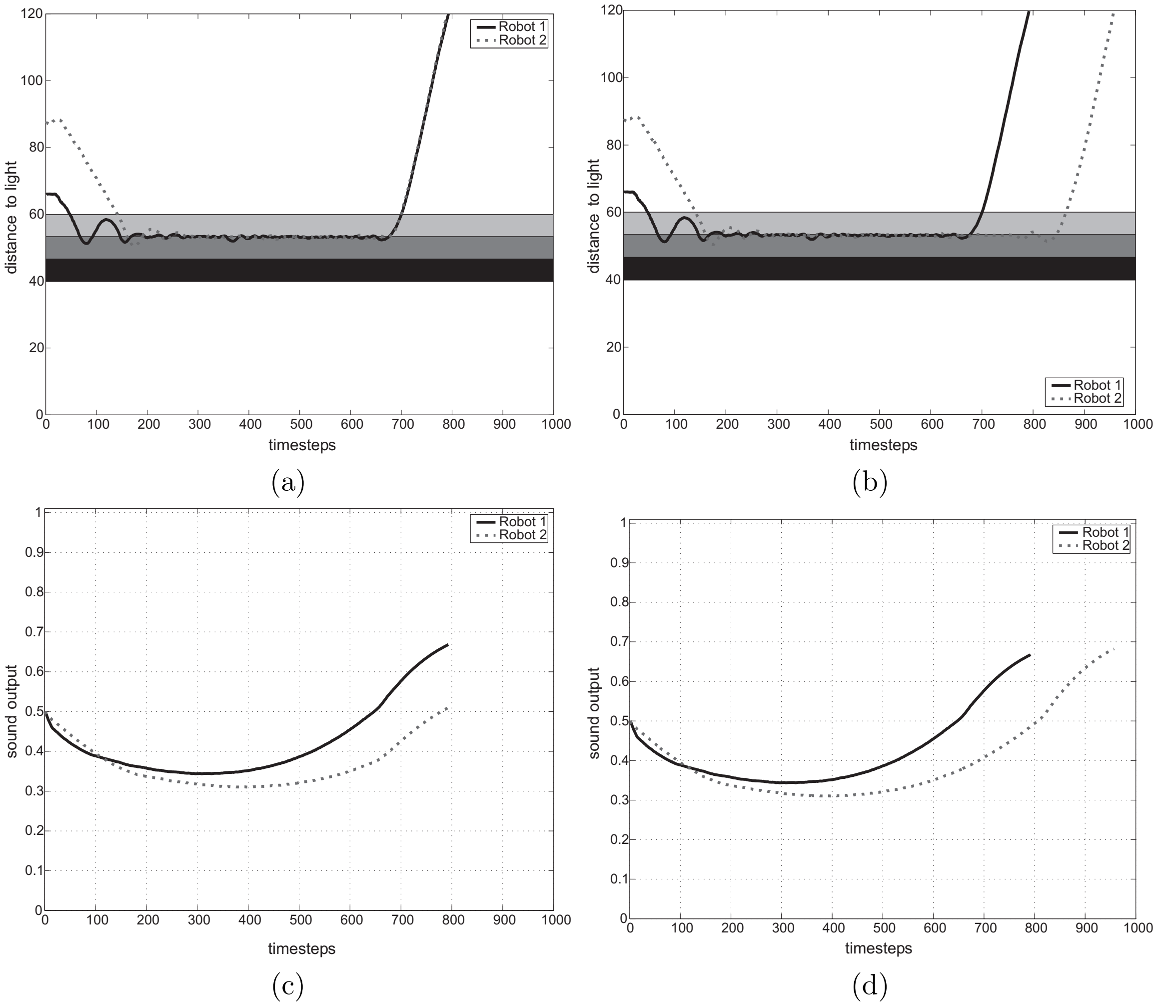
The results of post-evaluation analyses carried out so far have shown that in nine of the best evolved groups, the robots emit sound during the accomplishment of the task in *Env.B*. Note that the emission of sound is not demanded in order to navigate towards the target and discriminate *Env.A* from *Env.B*. Indeed, the task and the fitness function do not require the robots to display signalling behaviour (see section 4.3.4). Mechanisms for phototaxis, antiphototaxis, and memory are sufficient for each robot to accomplish the task. Therefore, in this section we show the results of further post-evaluation tests on those groups in which the robots emit sound during the accomplishment of the task. These tests aim to determine whether sound has a functional significance within the behavioural strategies of the groups and, if the answer is positive, to identify the adaptive function of sound use.

4.4.1.1 Behavioural features and mechanisms

We looked at the behaviour of the robots that emit sound during a successful trial in each type of environment. During each trial, we recorded for each robot of a group the distance to the light and the change over time of the sound output (i.e., cell potential of neuron N_{13} mapped into $[0.0, 1.0]$ by a sigmoid function σ). These two variables are recorded both in a *normal* condition and in a condition in which the robots can not hear each other's sound

group	*Env.A*				*Env.B*			
	fitness		*signalling (%)*		*fitness*		*signalling (%)*	
	mean	sd	mean	sd	mean	sd	mean	sd
g_1	1.92	0.31	0.00	0.00	1.98	0.13	17.39	0.30
g_2	1.94	0.28	0.72	3.72	1.99	0.00	18.22	1.36
g_5	1.99	0.10	0.00	0.00	1.98	0.10	13.36	1.58
g_6	1.96	0.21	0.00	0.00	1.99	0.11	16.47	2.38
g_7	1.99	0.11	0.00	0.00	1.95	0.21	15.06	2.82
g_8	1.96	0.25	0.00	0.00	1.99	0.02	16.47	2.08
g_9	1.99	0.12	0.00	0.00	1.97	0.16	16.38	2.62
g_{10}	1.91	0.31	0.00	0.00	1.91	0.36	0.00	0.00
g_{13}	1.87	0.43	1.72	8.14	1.95	0.09	20.88	2.44
g_{14}	1.96	0.17	0.00	0.00	1.98	0.17	0.00	0.00
g_{16}	1.89	0.33	0.00	0.00	1.94	0.27	0.00	0.00
g_{18}	1.81	0.45	0.00	0.00	1.87	0.16	0.00	0.00
g_{19}	1.91	0.27	0.00	0.00	1.98	0.06	12.65	0.99

Table 4.1: Results of post-evaluation tests showing for each best evolved successful group of each evolutionary run (g_i): the average and standard deviation of the fitness over 500 trials in *Env.A* (see columns 2, and 3) and in *Env.B* (see columns 6, and 7); the average and standard deviation of the percentage of timesteps sound was emitted by either robot over 500 trials in *Env.A* (see columns 4, and 5) and in *Env.B* (see columns 8, and 9).



(a) (b)

(c) (d)

Figure 4.4: The graphs show some features of the behaviour of the group of robots g_2 at each timestep of a successful trial in *Env.B*. Graphs (a) and (b) show the robots' distance to the light. The areas in shades of gray represent the circular band. Graphs (c) and (d) show the cell potential of neuron N_{13} mapped into $[0.0, 1.0]$ by a sigmoid function σ (i.e., the sound output) of each robot controller. Graphs (a) and (c) refer to the *normal* condition. Graphs (b) and (d) refer to the *not-other-sound* condition (i.e., the robots do not hear each other's sound). Robot 1—see continuous lines—is always initialised closer to the light than Robot 2—see dashed lines.

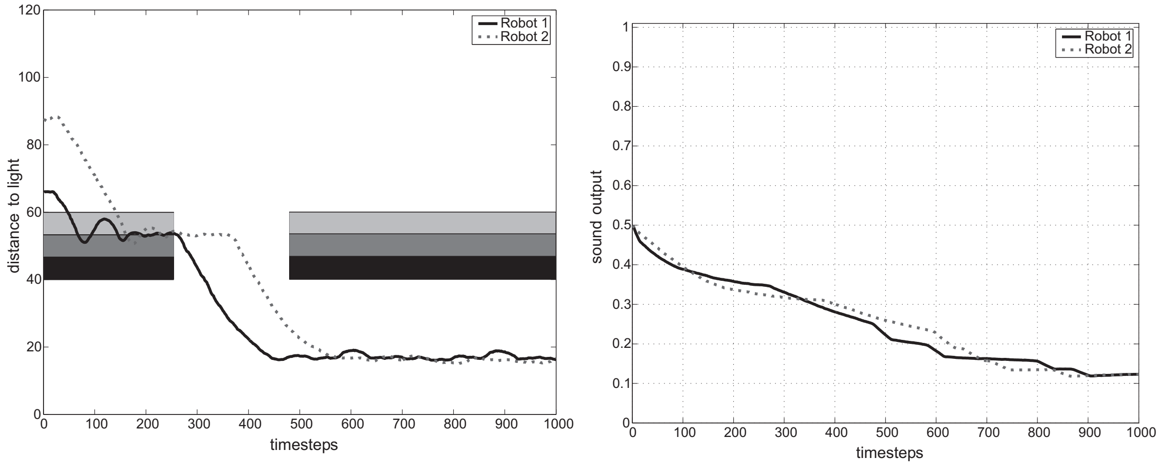
(i.e., the *not-other-sound* condition). In the latter circumstances, the input of neuron N_9 of each robot controller is set to 1 only if the sound in the environment is produced by the robot itself. Figure 4.4 shows the results of the tests for robots of group g_2 in *Env.B*. We show only the results of one signalling group (i.e., g_2) since it turned out that the groups that emit sound in *Env.B* share the same behavioural strategies. Therefore, everything that is said for group g_2 with respect to sound signalling, applies to groups g_1, g_5, g_6, g_7,

g_8, g_9, g_{13}, g_{19}.

In figure 4.4a, and 4.4b, continuous and dashed lines refer to the robot-light distances in, respectively, the *normal* condition and the *not-other-sound* condition. In both figures, the areas in shades of gray represent the circular band. From these figures, we can recognise three phases in the behaviour of the robots. In the first phase, the robot-light distance initially decreases for both robots (phototaxis phase). When the robots touch the band, the distance to the light remains quite constant as the robots circle around the band trying to find the *way in* zone (integration over time phase). In the third phase the robot-light distances increase and reach their maximum at the end of the trial (antiphototaxis phase). We immediately notice that the behaviour of the robots in the *normal* condition (see figure 4.4a) only slightly differs from what observed in the *not-other-sound* condition (see figure 4.4b). The only difference concerns the third phase. In particular, while in the *normal* condition both robots begin to move away from the light at the same time, in the *not-other-sound* condition Robot 2 initiates the antiphototactic behaviour after Robot 1. If observed with respect to how the robots' sound output unfolds in time, this small behavioural difference turns out to be an extremely indicative cue as to the function of sound.

Figure 4.4c, and 4.4d show that for both robots the sound output changes smoothly and in the same way in both conditions. During the phototaxis phase, the sound output decreases. During the integration over time phase, this trend is reversed. The sound output starts to increase up to the point at which its value rises over the threshold of 0.5. The increment seems to be induced by the persistence of a particular sensory state corresponding to the robot moving around the light on the band. Once the sound output of a robot increases over the threshold set to 0.5, that robot starts emitting a tone. In the *normal* condition we notice that, as soon as the sound output of Robot 1 rises over the threshold of 0.5 (see continuous line in figure 4.4c around timestep 650) both robots initiate an antiphototactic movement. Robot 2 leaves the band the moment Robot 1 emits a signal, despite the fact that its own sound output is not yet over the threshold of 0.5. Contrary to this, in the *not-other-sound* condition we notice that Robot 2 does not leave the band at the same time as Robot 1, but it initiates antiphototaxis only at the time when it starts emitting its own sound (see dashed line in figure 4.4d around timestep 830).

To qualitatively describe the behaviour of the robots in *Env.A*, we point the reader to figure 4.5; in figure 4.5a, continuous and dashed lines refer to the robot-light distances in the *normal* condition. From this figure, once again we can recognise three phases in the behaviour of the robots. In the first phase, the robot-light distance initially decreases for both robots (phototaxis phase). When the robots touch the band, the distance to the light remains quite constant as the robots circle around the band trying to find the *way in* zone (integration over time phase). In the third phase, the robot-light distances decrease as the robots find the *target area* after they approach the light through the *way in* zone. When signalling groups are located in *Env.A*, as can be observed in figure 4.5b, the robots' sound output undergoes a trend similar to the one shown in figure 4.4c. That is, it decreases during the initial phototactic phase and starts rising during the integration over time phase. However, when the robots are placed in *Env.A*, the increment of their sound output is interrupted by the encounter of the *way in* zone (see figure 4.5b, around



(a) (b)

Figure 4.5: The graphs show some features of the behaviour of the group of robots g_2 at each timestep of a successful trial in *Env.A*. Graph (a) shows the robots' distance to the light. The areas in shades of gray represent the circular band. Graph (c) shows the cell potential of neuron N_{13} mapped into $[0.0, 1.0]$ by a sigmoid function σ (i.e., the sound output) of each robot controller. Both graphs refer to the *normal* condition. Robot 1—see continuous lines—is always initialised closer to the light than Robot 2—see dashed lines.

280 timesteps). As soon as the robot gets closer to the light via the *way in* zone, the sound output begins to decrease. This process has been shaped by evolution in such a way that, in order for the sound output to rise over the threshold of 0.5, it must be the case that no *way in* zone has been encountered by the robots. In other words, it takes more or less the time to make a loop around the light while moving on the circular band for a robot's sound output to rise over the threshold. Consequently, when the robot is located in *Env.A*, no sound is emitted. Those post-evaluation trials in which sound has been recorded in *Env.A* in signalling groups (see table 4.1 columns 4, and 5, groups g_2, and g_{13}) were due to the effect of noise on perception and navigation trajectories; consequently, the sound output of either robot erroneously rose above the threshold.

4.4.1.2 The role of sound

The way in which the distance to the light and the sound output of each robot change over time in the two experimental conditions suggests that the sound is functionally relevant to the accomplishment of the task. In particular, the signalling behaviour seems to be strongly linked to mechanisms for environmental categorisation. As long as the latter mechanisms work properly, the emission of sound after approximately one loop around the light becomes a perceptual cue that reliably indicates to a robot the necessity to move away from the light. Moreover, sound has a communicative function: that is, once broadcast into the environment by one robot (e.g., Robot 1 in *normal* condition), it changes the behaviour of the other robot (i.e., Robot 2 in *normal* condition) which stops circuiting

Group g_2

Env.A				*Env.B*							
fitness		*signalling (%)*		*fitness*		*signalling (%)*		Robot 1 (d_f)		Robot 2 (d_f)	
mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
1.97	0.16	1.35	7.03	1.26	0.09	51.13	4.35	66.52	14.46	54.90	3.12
1.94	0.28	0.72	3.72	1.99	0.00	18.22	1.36	119.65	0.20	119.64	0.20

Table 4.2: *Deaf* setup (robots' sound inputs set to 0): results of post-evaluation test showing for group g_2 the average and standard deviation of the fitness over 500 trials in *Env.A* (see columns 1, and 2) and in *Env.B* (see columns 5, and 6); the average and standard deviation of the percentage of timesteps the sound was on by either robot over 500 trials in *Env.A* (see columns 3, and 4) and in *Env.B* (see columns 7, and 8); the average and standard deviation of the final distance (d_f) of each robot to the light in *Env.B* (see columns 9, 10, 11, and 12). The row in gray shows again the result of group g_2 in the normal condition, with no disruptions applied to the propagation of sound signals.

around the light and initiates antiphototaxis (see figure 4.4a and 4.4b).

To further test the causal relationship between the emission of sound and the switch from phototaxis to antiphototaxis, we performed further post-evaluation tests. In these tests, we post-evaluated group g_2 for 500 trials in *Env.A* and 500 trials in *Env.B*, in conditions in which the robots are not capable of perceiving sound. That is, their sound input is set to 0 regardless of whether any agent emits a signal. We refer to this condition as the *deaf* setup. We remind the reader that similar phenomena to the one concerning g_2 and illustrated in table 4.2, have been observed for all the other signalling groups. As far as it concerns *Env.A*, the average fitness of the group does not differ much from the average fitness obtained in the *normal* setup (see table 4.2 column 1 and 2). Concerning *Env.B*, the average fitness of the group is lower than the average fitness recorded in the *normal* setup (see table 4.2 column 5, and 6). Moreover, the robots' average final distance to the light is only about the same as the radius of the outer edge of the band (i.e., 60 cm to the light; see table 4.2 columns 9, 10, 11, and 12). Given that the robots never collided, the decrease of the average fitness recorded in *Env.B* in the *deaf* setup can only be attributed to the fact that the robots do not perform antiphototaxis. This confirms that, in conditions in which the robots can not hear any sound, they do not switch from phototaxis to antiphototaxis. The role of sound is indeed to trigger antiphototaxis in both the emitter and the robot that is not emitting a tone yet.

4.4.1.3 Neural functionality and lesion analysis

In the case of a robot of a signalling group (e.g., g_2), we have seen above that the output of N_{13} (sound output) is integrating the information over time by rising and passing over the threshold of 0.5 when the discrimination is performed. But since this is not the case for non-signalling groups (e.g., g_{10}), we should look elsewhere to find a neuron that performs the integration. In figure 4.6 we plot the distance to the light as well as the firing rates of all the neurons of the network, for one robot of groups g_2 and g_{10}, over time for both environments, during a successful trial. Notice that these values are not passed through the sigmoid function, that is why the plot for N_{13} is different than the one in figure 4.4c, where we plot the sound output (cell potential passed through the sigmoid function). We notice that for group g_2 there is N_3 (that takes input from one floor sensor) whose firing rate's evolution through time has the characteristics defined above: starts rising (from 0) when the robots start circling around the band (approximately constant distance to light) and, in case the *way in* zone is encountered, it stops rising and starts decreasing; in case the agent is in *Env.B*, it continues to rise and eventually passes beyond the threshold. Immediately after we notice that the robot is leaving the band. This suggests that there is also an internal integration mechanism, apart from the one present in the activation of the sound output neuron. Concerning the behaviour of the robot of group g_{10}, the neuron that plays this role is N_5 (that takes input from a combination of proximity sensors).

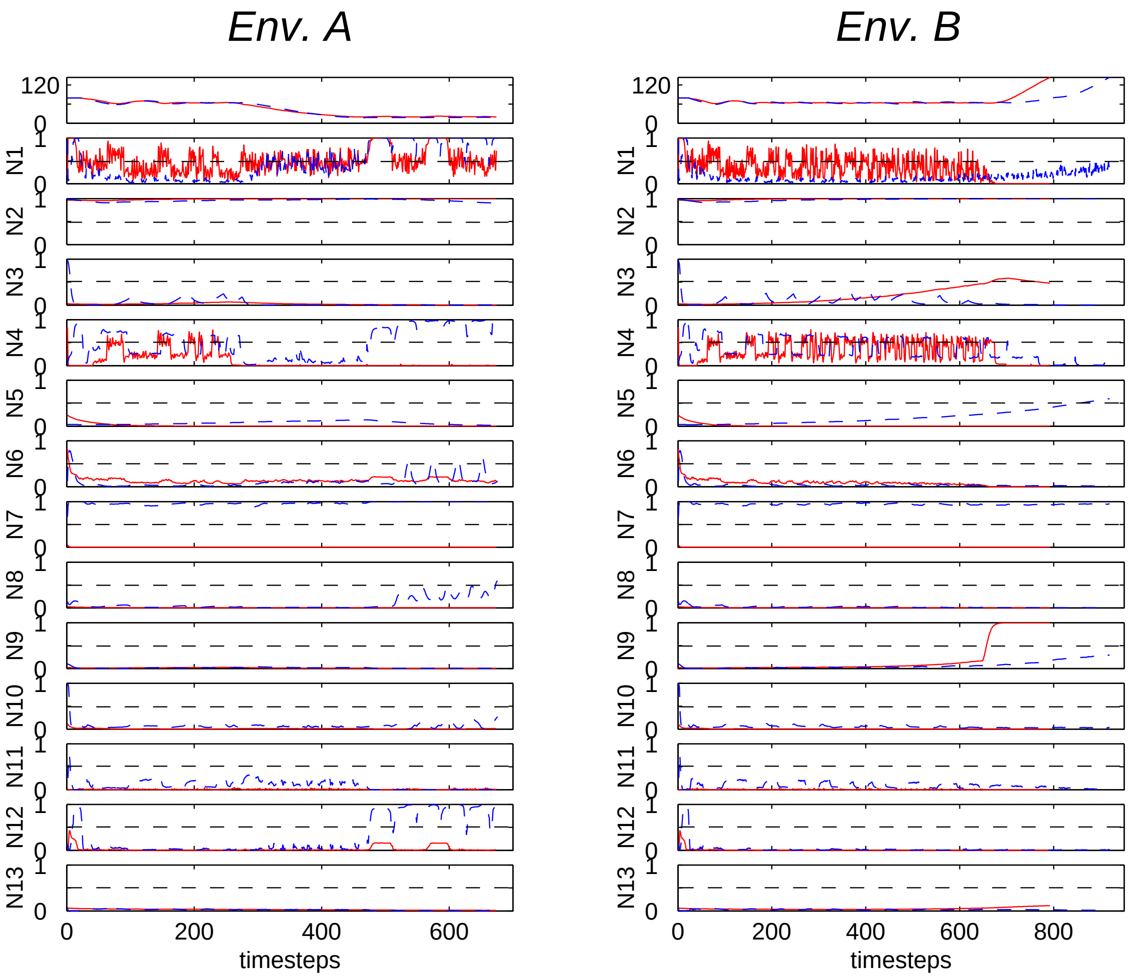


Figure 4.6: Plot of the distance to the light (box at the top) and of the firing rate of all neurons with time for a robot of g_2 (continuous lines) and g_{10} (dashed lines)

Table 4.3: Results of the lesion analysis for groups g_2 and g_{10}. After introducing a lesion in all neurons of the network ($N_1 - N_{13}$), we measure the mean of the average final distance d_f of the two robots in 500 evaluations in both environments.

neuron	mean of the average d_f of the two robots			
	g_{10}		g_2	
	Env.A	*Env.B*	*Env.A*	*Env.B*
N_1	77.25	92.09	63.73	62.49
N_2	21.26	118.11	18.08	119.64
N_3	17.72	78.56	17.37	53.23
N_4	45.94	45.15	38.93	41.95
N_5	16.40	53.07	18.09	119.66
N_6	20.18	118.82	17.69	119.47
N_7	24.31	119.63	18.83	119.65
N_8	22.02	119.68	17.86	119.63
N_9	17.15	60.57	16.72	53.13
N_{10}	19.65	117.37	17.51	119.65
N_{11}	108.99	112.57	102.38	106.88
N_{12}	81.34	90.21	42.58	43.98
N_{13}	20.09	119.59	17.87	58.58

To prove that these neurons are indeed essential for the discrimination to take place, we introduce lesions in the robot controllers which selectively damage the functionality of one neuron at a time.[4] Specifically, we confine the value of the firing rate of a neuron to the average value observed throughout the robots lifetime and we re-evaluate the system with the new conditions. We are interested in discovering the functionality of each neuron in the network and also which neurons are indeed tied to the discrimination mechanism. Table 4.3 gives us the results of this lesion analysis, by displaying the average final distance of the two agents over 500 evaluations in *Env.A* and 500 in *Env.B*, for g_{10} and g_2. What we notice is that N_3 is the integration neuron for g_2 and N_5 for g_{10}, and without them the robots are unable to trigger antiphototaxis (the average distances for *Env.A* suggest that they do find the *way in* zone, while for *Env.B* that they stay on the band). For g_2, the sound input and output neurons are essential to complete the task in *Env.B.* What is surprising though is that if N_9 (sound input neuron) is disturbed, g_{10}'s performance gets disrupted, despite the fact that as we saw this genotype does not rely on the presence of sound to trigger antiphototaxis. In detail, what happens is that the disruption prevents robots from performing antiphototaxis in *Env.B*, while they are able to find the *way in* zone in *Env.A*. In section 4.7 we provide further details on this behaviour.

To summarise, we can say that for all the best evolved groups of robots, there is a

---

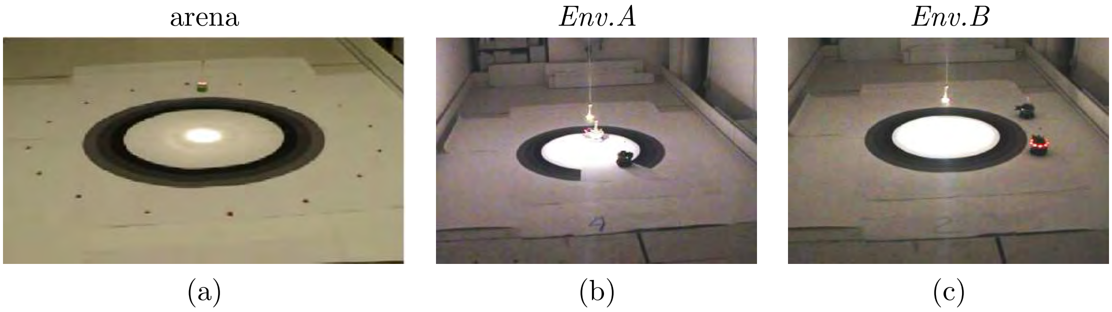
[4] A more complete lesion analysis should also investigate the effect of introduction of lesions to groups of neurons, in order to test the relationship and interplay among neurons. Instead, our analysis is simplified and focuses on the study of the functionality of one neuron at a time, which proved to be enough to infer some basic operational principles of our system.

neuron other than the sound output neuron whose firing rate behaves similarly to neuron N_{13} of the robots in group g_2. That is, there is a neuron whose firing rate increases in response to the persistence of the sensory states associated with moving around the light on the band. For groups that never emit sound (i.e., g_{10}, g_{14}, g_{16}, g_{18}), if this increase is not interrupted by the encounter of the *way in* zone, it eventually induces antiphototaxis. For groups that emit sound (i.e., g_1, g_2, g_5, g_6, g_7, g_8, g_9, g_{13}, g_{19}), this mechanism is linked to the behaviour of neuron N_{13} as shown in figure 4.4c. The relationship between mechanisms for integration of time and neuron N_{13} (which is responsible for sound emission) is the basic difference between signalling and non-signalling groups.

4.5 Transfer to real robots

The task described in this chapter is characterised by the fact that not only the change but also the persistence of particular sensorial states is directly linked to the effectiveness of the evolved strategies. These strategies are generated by robot controllers developed in a simulated world that is responsible for modelling the sensory states of *s-bots* acting in *Env.A* or *Env.B*. Our simulated world (see section 4.3.2) models only a small subset of the *s-bot* world physics, since it has been designed to speed up a particularly long evaluation process (i.e., 12,000 generations, 100 genotypes, 10 evaluation trials for each genotype, 1,000 simulated time cycles for each trial). Also, it does not model potential inter-robot differences. As mentioned in section 4.3.2, we compensate for the effect of those physical phenomena not modelled (e.g., acceleration, friction, inter-robot differences, etc.), by adding random noise to the light and proximity sensor readings, the motor outputs, the position of the robot, and the reading of the two floor sensors. However, there is always the risk that the physics of our simulated world are insufficiently or incorrectly defined, and that the evolved behavioural strategies exploit loopholes which limit their effectiveness to an unrealistic scenario. Porting the controllers evolved in simulation onto a real robot is the best way to rule out the above mentioned problem (Brooks, 1992). As pointed out in section 4.2, this step has not previously been taken in previous research work in which CTRNNs have been evolved to deal with tasks that required the integration over time of sensory states. In this section, we provide evidence of the "portability" of the evolved controllers by showing the results of tests in which real robots are repeatedly evaluated in *Env.A* and *Env.B*. We chose to re-evaluate the controller of the successful group g_2 because this group during post-evaluation achieved a very high performance, but also because in preliminary tests, among other equally successful controllers, this one seemed to achieve the best sensory-motor coordination when downloaded on a group of real robots. Experiments are performed with groups of two and four *s-bots*.

Jakobi (1997) claims that the robot does not have to act identically in simulation and reality for the porting to be called successful. In fact, it is enough that its behaviour satisfies some criteria defined by the experimenter. Following this principle, real robots are considered successful if they carry out the main requirements of our task. That is, the robots have to reach the band in shades of gray regardless of the type of environment and subsequently (i) end up in the *target area* in *Env.A*, without crossing the inner black edge



(a) (b) (c)

Figure 4.7: The experimental setup. (a) A picture of the arena, with the points around the band showing the locations where the robots were randomly positioned. (b) A snapshot of a trial in which two robots find the *way in* zone in *Env.A*. (c) A snapshot of a trial in *Env.B*. The robot with the lighter turret colour is the one that has signalled the absence of a *way in* zone. Both robots have left the band and are performing antiphototaxis.

of the circular band; (ii) end up as far as possible from the light in *Env.B*. The robots should also avoid collisions.

4.5.1 Experiments with two *s-bots*

In our real-world experimental setup, two *s-bots* ($s\text{-}bot_1$ and $s\text{-}bot_2$) are randomly positioned at a distance of 85 cm from the light. We performed 40 trials, 20 in each environment. Each trial differs from the others for the randomly defined initial position and orientation of the robots, and for the position of the *way in* zone in *Env.A*. The initial position of the robots is randomly chosen among one of the sixteen possible starting positions which surround the light (see figure 4.7a). The width of the *way in* zone is fixed to 45°, which is the smallest value encountered during evolution and the most difficult case for a possible misidentification of an *Env.A* for an *Env.B*. The *s-bots* proved to be 100% successful in both environments: there were no mistakes in discrimination, no collisions, and no crossing of the black edge of the band.[5] As was the case for the simulated robots of group g_2, the *s-bots* accomplished the task by using sound in a communicative context. That is, the sound emitted by one *s-bot* triggers antiphototaxis in both robots. The following paragraphs provide further quantitative descriptions of the behaviour of simulated and real robots. This data will help to quantify the extent to which the behaviour of simulated robots diverges from the behaviour of real robots and to evaluate the reliability of our simulated world as a tool for developing controllers for real robots.

Given the nature of the successful strategy of group g_2, the start of the emission of a tone can be used as a sign which precisely indicates when an *s-bot* has reached the conclusion that it is located in *Env.B* rather than *Env.A*. We compute the offset between the entrance position in the circular band of the robot that first emits a signal and the position at which this robot starts to signal. This measure, called offset Δ, takes value

---

[5] The movies that correspond to these experiments can be found at `http://iridia.ulb.ac.be/supp/IridiaSupp2006-007`.

0° if the robot signals exactly after covering a complete loop around the circular band. Negative values of the offset Δ suggest that the robot signals before having performed a complete loop, while positive values correspond to the situation in which the robot emits a tone after having performed a loop around the light. The offset Δ is used to compare the behaviour of simulated and real robots, and is computed as follows (Tuci et al., 2004):

$$\begin{aligned} \Delta &= |\alpha(t_e, t_s)| - 2\pi, && (4.2)\\ \alpha(t_1, t_2) &= \sum_{t=t_1}^{t_2-1} \widehat{\mathbf{AOB}}, \qquad \mathbf{A} = \mathbf{X}_t, \mathbf{B} = \mathbf{X}_{t+1} && (4.3) \end{aligned}$$

where $\mathbf{O}$ corresponds to the position of the light, and α is the angular displacement of the robot around the light from the starting position (the position at time t_e when the robot touches the circular band) to the signalling position (the position at time t_s when the robot starts signalling). α is computed summing up all the convex angles $\widehat{\mathbf{AOB}}$ comprised between two consecutive positions of the robot $\mathbf{A}$ and $\mathbf{B}$.

During the tests on real robots, we observed that in *Env.B* it is always *s-bot*$_1$ that emits a signal. As shown in table 4.4, we see that the *s-bot* that first emits a signal does so on average before completing a loop. However, given that the value of the offset Δ is larger than $-45°$ (the width of the *way in* zone is equal to 45°), the group does not run into the risk of misinterpreting an *Env.A* as an *Env.B*. Further tests have proved that, if left to act alone in an *Env.B*, *s-bot*$_2$ always signals after completing a loop (i.e., positive offset Δ, data not shown). This result can be accounted for by noting the existence of various arbitrary mechanical and sensor differences between the two *s-bots*; inter-robot differences that are impractical to include in the simulated world. Contrary to the *s-bots*, the simulated robots of group g_2 signal on average after completing the loop (see table 4.4). The mismatch between the behaviour of simulated and real robots controlled by the same neural network is an estimate of the magnitude of the divergence between the simulated and real worlds. However, given that our real robots were 100% successful in both environments, we conclude that the noise injected into the simulated world was sufficient to cross the "reality gap" (Jakobi, 1997) and to capture the variability of the behaviour of sensors and actuators of real hardware which can easily disrupt the effectiveness of the evolved neural mechanisms. Note that the successful porting of the controller of a group (i.e., g_2) does not necessarily imply that controllers of other groups that were successful in simulation would be equally successful in guiding real robots. For example, if the effects of inter-robot differences on the mechanisms used for environmental discrimination induce robots to anticipate (with respect to what the group does in simulation) the emission of a signal, then simulated groups with a $\Delta \in [-45°, 0°)$ could be more likely to fail. In fact, these groups might fall into the error of signalling and consequently performing antiphototaxis even if placed in an *Env.A*.

4.5.2 Experiments with four *s-bots*

We also performed a further experiment by porting the controllers of group g_2 to a group of four different *s-bots*—namely *s-bot*$_3$, *s-bot*$_4$, *s-bot*$_5$ and *s-bot*$_6$. The aim of these tests is

groups	Offset Δ	
	avg (degree)	sd (degree)
two *s-bots*	-30.6	11.75
four *s-bots*	+18.22	12.97
simulated robots	+31.6	16.05

Table 4.4: Average and standard deviation of the offset Δ recorded for different group types.

twofold: first, we test the ability of the evolved controllers to accomplish the task despite the cardinality of the group being higher than that experienced during the evolutionary phase. Second, we evaluate the effectiveness of our controllers with respect to individual differences among the robots. The experiment consists of evaluating for 10 trials the four-robots group in each environment.

The results are once again almost perfect.[6] In all trials in *Env.A*, the robots found the *way in* zone without erroneously emitting a tone or crossing the inner black edge of the band in shades of gray. In *Env.B* we noticed that *s-bot*$_3$ was always signalling first, but never too early. In all trials, *s-bot*$_5$ never emitted a tone and in a separate test we discovered that it was signalling much too late, even after more than two full loops around the band.[7] Nevertheless, all robots reacted properly to the signal emitted by *s-bot*$_3$, left the band, and reached the appropriate distance from the light. In table 4.4 we see that, with respect to the offset Δ, the behaviour of the four-robots group is closer to simulation than the behaviour of the two-robots group. This is because the robot to signal first in the four-robot group (*s-bot*$_3$) does so with an offset Δ closer to the one of simulated robots than the robot to signal first in the two-robot experiment (*s-bot*$_1$).

In one trial, *s-bot*$_5$ while performing antiphototaxis as a reaction to the sound emitted by *s-bot*$_3$, made a turn of 180° and started moving wrongly towards the light. In all other trials though we did not observe this error and it looks to be a hardware crash. Another error which was not expected and revealed some property of our controller about which we would not have found out had we not performed the four-robot test, is the fact that the robot-robot avoidance behaviour does not work while the robots perform antiphototaxis. In fact, as they leave the band after they perceive a sound signal, they ignore the information coming from the proximity sensors, with the consequence that, should they encounter another agent on their way, they collide against each other. A possible explanation for this is that this condition was never encountered during evolution, and therefore the mechanism shaped was confined to just leaving the band without paying attention to obstacles (other robots). Finally, by allowing more agents to interact in the target area, we discovered that the robot-robot avoidance mechanism is different once the robots are interacting there—a case not often encountered during the two robot experiments–and

---

[6] The movies that correspond to these experiments can be found at `http://iridia.ulb.ac.be/supp/IridiaSupp2006-007`.

[7] The light sensors of *s-bot*$_5$ are the reason for this behaviour. In fact, their readings proved to be different from those of the other robots (data not shown). By comparing the behaviour of this robot with *s-bot*$_1$, one can get an idea of the magnitude of the inter-robot differences.

can be described as follows: if an agent detects others in its vicinity, it stops and spins until the other agents have moved away.

4.6 On the adaptive significance of signalling

The results illustrated in section 4.4 have shown that the majority of the successful strategies employ signalling behaviour and communication among the members of the groups. This suggests that our decision to equip the robots with "ears" and a "mouth" turned out to be helpful. However, by simply looking at the characteristics of our model, we cannot necessarily see why evolution exploited these robots' structures to develop a simple form of communication. In principle, groups in which the use of sound is functionally relevant for the success of the group, and groups in which it is not, can be equally successful. Yet, the majority of the evolutionary runs that ended successfully (i.e., nine out of thirteen best evolved groups) are characterised by group strategies that make use of sound signalling and communication among the robots (see table 4.1). How can we account for this result?

It might be that there is in fact no selective advantage for groups in which the use of sound is functionally relevant to their success with respect to alternative group types. The evolution of signalling might simply be due to the effect of statistical drift of genetic material over time in populations of simulated agents (i.e., genetic drift). However, we collected evidence that rules out the genetic drift hypothesis, and that supports the idea that there are selective pressures which favour signalling over non-signalling groups. The rest of the section is dedicated to this issue.

4.6.1 Functions of sound signalling

We started our analysis by trying to understand whether during evolution sound had fulfilled functions other than the one we observed in the best evolved groups of robots during the post-evaluation tests shown in section 4.4.1. To do this, we post-evaluated (500 times in each type of environment) all the best groups at each generation (1 to 12,000) of all the successful evolutionary runs. During this post-evaluation, we recorded the average fitness in each environment and the average percentage of time per environment either robot emits a signal during a trial. After post-evaluating these groups, we isolated those whose average fitness was higher than 1.8. We noticed that after having excluded (i) those groups that signal throughout the entire duration of a trial in both environments,[8] (ii) those groups that never signal in a trial in both environments, and (iii) those groups in which sound was not functionally relevant for their behavioural strategies, we were left with groups that signal only in *Env.B* for an average time of about one fourth of the duration of a trial. Further investigation on the behaviour of these groups revealed that in all of them sound was fulfilling one and only one function: triggering antiphototaxis in *Env.B*.

---

[8] We do not further analyse the cases in which the robots signal throughout the entire duration of a trial since we consider it obvious that in these cases the sound-emitting behaviour does not serve any specific function.

In other words, looking at the behaviour of all successful signalling groups of any evolutionary simulation run, we discovered that whenever signalling is functionally relevant to the success of the group, it is employed by the robots in *Env.B* as a self-produced perceptual cue. This cue induces the emitter as well as the other robot of the group to change its behaviour from light-seeking to light-avoidance. This evidence constrains our investigation on the adaptive significance of sound signalling to only a specific case in which we can arbitrarily associate to sound two functionalities: on the one hand, sound is the means by which a robot emitter switches from phototaxis to antiphototaxis. We refer to this as the *solitary* function. On the other hand, sound is the means by which the robot emitter influences the behaviour of the other robot. In fact, the perception of sound triggers antiphototaxis in the emitter as well as in the robot that is not yet emitting a tone (see figure 4.4a and 4.4c). We refer to this as the *social* function. In the following, we illustrate the results of post-evaluations that prove and explain why it is the latter functionality which makes a group of signalling robots better adapted than other group types.

4.6.2 The social function of sound signalling as a means to obtain robustness

The statistics shown in table 4.5 refer to a series of tests in which we post-evaluated (500 times in each environment) 100 different groups of robots of five different evolutionary runs (runs 2, 10, 14, 16, 18), chosen among the best of each generation whose average fitness was higher than 1.8. As far as it concerns run 2, we post evaluated: (i) 100 groups that use sound signalling in the *normal* setup (see table 4.5 second row "sig") and in the *not-other-sound* setup (see table 4.5 fourth row "not-other"); (ii) 100 groups that do not use sound signalling (see table 4.5 third row "non-sig"). Recall that the *not-other-sound* setup refers to the case in which the robots do not hear each other's sound (see also section 4.4.1). The 100 non-signalling groups of robots of evolutionary run 2 are "predecessors" of the signalling one. That is, they were the best groups some generations before the evolution of successful signalling groups.

By looking at the statistics shown in table 4.5 we notice that: (a) the mean fitness of signalling groups (run 2) is significantly higher than the fitness of any of the non-signalling groups (run 2 "not-sig", 10, 14, 16, and 18, pairwise Wilcoxon test with 99 % confidence interval), for both environments; (b) the standard deviation of the fitness of signalling groups (run 2) is smaller than the standard deviation of the fitness of any of the non-signalling groups for both environments (run 2 "not-sig", 10, 14, 16, and 18), for both environments; (c) in *Env.B*, the mean fitness of signalling groups (run 2) recorded in the *not-other-sound* condition is significantly smaller than the mean fitness of any of the non-signalling groups (run 2 "not-sig", 10, 14, 16, and 18, pairwise Wilcoxon test with 99% confidence interval). We consider (a) and (b) empirical evidence which suggests that indeed signalling groups are on average better than non-signalling groups. Notice that, although the difference among the groups is small, during evolution it may have influenced the distribution of genetic material and consequently the emergence of the behavioural strategies. For the sake of completeness, we also show the lower and upper quartile and the median of the distributions. This data confirms that the difference in performance between the two groups seems to lie in the fact that non-signalling groups display a slightly worse performance than signalling groups in few cases (see lower quartiles, run 2 "sig" and "non-sig"). We consider (c) evidence suggesting that the beneficial effect of signalling is not linked to the solitary function, since if we prevent signalling robots from hearing each other's sound (i.e., the *not-other-sound* setup) the solitary function is not by itself sufficient to make the robots on average better than those that do not use signalling at all. Consequently, it appears that groups of robots that use sound signalling have a selective advantage over other types of groups, due to the social function of signalling.

In particular, we believe that the selective advantage of signalling groups is given by the beneficial effects of communication with respect to a robust disambiguation of *Env.A* from *Env.B*. The beneficial effect corresponds to robust individual decision-making and faster group reaction, since signaller and hearer react at the same time. Moreover, the effectiveness of the mechanisms which integrate sensory information over time in order to produce the categorisation of the environment is disrupted by the random noise explicitly injected into the simulated world, which strongly affects the sensors' reading and the outcome of any "planned" action. However, by communicating the outcome of their decision about the state of the environment, signalling groups, contrary to other types of group, might exploit social behaviour to counterbalance the disruptive effect of noise on individual mechanisms for environmental discrimination. In total, in those groups in

	run	groups	mean	sd	lower quartile	median	upper quartile
Env.A	2	sig	1.982	0.151	2	2	2
		non-sig	1.942	0.251	2	2	2
		not-other	1.870	0.391	2	2	2
	10	non-sig	1.874	0.380	2	2	2
	14	non-sig	1.897	0.331	2	2	2
	16	non-sig	1.749	0.495	2	2	2
	18	non-sig	1.871	0.402	2	2	2
Env.B	2	sig	1.989	0.082	1.995	1.996	1.997
		non-sig	1.923	0.261	1.964	1.995	1.997
		not-other	1.747	0.268	1.589	1.760	1.982
	10	non-sig	1.905	0.308	1.966	1.995	1.997
	14	non-sig	1.943	0.226	1.993	1.996	1.997
	16	non-sig	1.945	0.210	1.992	1.995	1.997
	18	non-sig	1.880	0.326	1.918	1.995	1.997

Table 4.5: The table shows the statistics of post-evaluation tests in which 100 different groups of robots of five different evolutionary runs (runs 2, 10, 14, 16, 18), chosen among the best of each generation whose average fitness was higher than 1.8. For run 2, we post evaluated: (i) 100 groups that use sound signalling in the *normal* setup (see row "sig") and in the *not-other-sound* setup (see row "not-other"); (ii) 100 groups that do not use sound signalling (see row "non-sig").

which antiphototaxis is triggered by the perception of sound rather than by an internal state of the controller, a robot which by itself is not capable or not ready yet to make a decision concerning the nature of the environment can rely on the decision taken by the other robot of the group. Therefore, by reacting to the sound signal emitted by the group mate, a robot initiates an action (i.e., antiphototaxis) which it may not have been capable of, or ready to perform, otherwise.

The experiments performed on real hardware provide perfect examples of the benefits of communication and social behaviour, given the presence of severe disruptions due to inter-robot differences. For example, in the experiments with two real robots, we have seen that *s-bot*$_2$ signals always later than *s-bot*$_1$ (see section 4.5.1). In the four-robots experiments (see section 4.5.2), we noticed that *s-bot*$_4$, *s-bot*$_5$ and *s-bot*$_6$ repeatedly benefit from the sound signal emitted by *s-bot*$_3$, which is the fastest—as well as proving extremely accurate—robot to signal the absence of the *way in* zone in *Env.B*. If a robot that reacts to the non-self produced sound could not have exploited the signal emitted by the other member of its group, it would have wasted precious time orbiting around the light. Eventually, it would have switched to antiphototactic behaviour, but due to time limits it would not have been able to reach the maximum possible distance to the light (see d_f in section 4.3.4). Consequently, the fitness of the group would have been lower.

The performance of signalling groups not only exceeds the performance of non-signalling groups in *Env.B*, but also in *Env.A* (pairwise Wilcoxon test with a 99% confidence interval). It seems that signalling groups are better adapted to the “danger” of discrimination mistakes in *Env.A* than are non-signalling groups, and thus “early” signalling seems to be an issue that has been taken care of by evolution. Our speculation is that once signalling groups evolve, their signalling behaviour is *refined*, by categorising the world later than in the case of non-signalling groups. Indeed, in table 4.6 we can clearly see that signalling groups tend to initiate antiphototaxis later than non-signalling groups. This happens in order to ensure that the chances of a potential disadvantage resulting from social behaviour are minimised (e.g., see table 4.6: simulated robots of group g_2 signal on average after completing a loop—rather late). In other words, the use of communication in a system can also affect aspects of the behaviour not directly related to communication (i.e., the process of integration of inputs over time). This explains the low performance recorded in the *not-other-sound* condition, compared to the *normal* condition. When robots emit signals later (high offset Δ), the system becomes more robust because the risk of a discrimination mistake in *Env.A* is minimised, at the cost of triggering antiphototaxis in *Env.B* somewhat later. However, this is counterbalanced by the effect of the social behaviour as explained above. To summarise, communication delays the moment of categorisation (larger offset Δ), and at the same time anticipates the collective response: putting robustness in *Env.A* and social behaviour in *Env.B* together, we can account for the selective advantage of communication.

4.6.3 A further set of evolutionary simulations

The analysis detailed in section 4.6 suggests that the selective advantage of signalling over non-signalling groups is the reason why we observe the evolution of signalling groups.

Offset Δ		
group	mean	sd
g_1	+54.1	20.23
g_2	+31.60	16.05
g_5	-14.35	23.80
g_6	+7.51	22.63
g_7	+89.23	44.40
g_8	+25.12	31.12
g_9	+76.33	42.24
g_{10}	+10.01	24.63
g_{13}	+61.12	29.30
g_{14}	-8.31	22.25
g_{16}	+6.95	35.74
g_{18}	+20.89	19.42
g_{19}	-8.08	16.24

Table 4.6: Mean and standard deviation of the offset Δ recorded for different groups, in 1,000 trials in *Env.B*. Rows in gray correspond to non-signalling groups, rows in white to signalling groups.

Moreover, it suggests that it is the social function of signalling (the communication resulting from it) that makes these groups more fit than others. In other words, we can attribute the evolution of signalling to its social function, and thus to the effect of emitted signals on other members of the group. To further clarify this issue and to provide compelling evidence about it, we decided to run further evolutionary simulations. In these simulations we introduced a simple modification to the characteristics of the robot's controller. This modification is meant to remove those circumstances which presumably made it harder for evolution to assemble the neural mechanisms in which the relationship between the perception of sound and antiphototaxis takes only place in a robot that is currently emitting a tone. More specifically, we explicitly demand that the solitary and the social function of acoustic signals are disentangled.

In particular, in this further set of evolutionary runs, the robots are controlled by the same type of thirteen neuron CTRNNs as described in section 4.3.2. However, contrary to the evolutionary runs illustrated in section 4.4, in this set of simulations, neuron N_{10} of each robot controller receives a binary input which is set to 1 when the other robot of the group is emitting a tone and to 0 otherwise, while neuron N_9 receives a binary input which is set to 1 only when the robot itself is emitting a tone and to 0 otherwise. This modification is introduced to allow the robot to distinguish between self (i.e., the input to neuron N_9) and non-self (i.e., the input to neuron N_{10}) produced sound. We assume that such distinction can be performed by real robots, for example by exploiting frequency differences among the signals—which is feasible on the *s-bot*, at least for two different frequencies. However, it should be stressed that such a system is non-scalable as it demand the dedication of a frequency to each robot. However, the number of frequencies that can be reliably discriminated is very limited and currently no more than two.[9]

This setup should make equally possible the evolution of groups of robots which discriminate between *Env.A* and *Env.B* through the following processes:

- *B1*: processes that do not rely on the perception of sound signals (as in non-signalling groups)
- *B2*: processes only relying on the perception of self-produced sound (thus not making use of communication, but making use of sound signalling)
- *B3*: processes only relying on the perception of non-self-produced sound (thus relying only on the communication resulting from the sound signalling)
- *B4*: processes that rely on the perception of any sound signal, either self or non-self produced (thus relying on the use of sound signalling, regardless of its function—solitary or social)

We run a new set of twenty evolutionary simulations and subsequently we post-evaluate the genotypes of each evolutionary run that achieved fitness over 1.8 during evolution. The groups of robots controlled by networks built from the best evolved genotypes are referred to as $g{*}_1$-$g{*}_{20}$, respectively. The results of the post-evaluation phase are shown in table 4.7. We notice that fourteen out of twenty best evolved genotypes obtained an average fitness higher than 1.8 in both environments. Among the successful groups, nine ($g{*}_1$, $g{*}_7$, $g{*}_{10}$, $g{*}_{13}$, $g{*}_{15}$, $g{*}_{16}$, $g{*}_{17}$, $g{*}_{18}$ and $g{*}_{19}$) make use of signalling (see table 4.7 columns 4, 5, 8, and 9, which refer to the average percentage and standard deviation of the time either robot emits a signal during a trial). The remaining four groups use process *B1* to perform the discrimination, that is they do not rely on sound signalling to solve the task.

To unveil the relationship between the emission of sound signals and the completion of the task we perform further post-evaluation tests (i.e., 500 trials in *Env.A* and 500 trials in *Env.B*), on the successful signalling groups in three different conditions: (i) the *deaf* setup with both sounds inputs set to 0; (ii) the *not-self-sound* setup in which the robots can only perceive the non-self produced sound. The input corresponding to the self produced sound is set to 0; (iii) the *not-other-sound* setup in which the robots can only perceive the self produced sound. The input corresponding to the non-self produced sound is set to 0. The results of these analyses are shown in table 4.8.

The performance of groups $g{*}_1$, $g{*}_{10}$, $g{*}_{15}$, $g{*}_{17}$ and $g{*}_{18}$ is disrupted only in *Env.B*, and only in the *deaf* and *non-other-sound* setup (see table 4.8 columns 4, and 12). Thus, we can conclude that these groups (i) do not use sound in *Env.A*; (ii) use the non-self produced sound to trigger antiphototaxis in *Env.B*. Concerning groups $g{*}_{13}$, $g{*}_{16}$ and $g{*}_{19}$, their performance is mainly disrupted in *Env.A*, and only in the *deaf* and *not-self-sound* setup (see table 4.8 columns 2, and 10). Through an analysis of the behaviour of

---

[9] Notice that these experiments were performed in order to enforce the hypothesis formulated above; the fact that we use experimenter-introduced ways for discrimination between self and non-self sound components is solely due to the fact that it helps us test that hypothesis. Otherwise, we believe that such a methodology might bear upon the scalability of the system and should normally be avoided when designing communication systems for robots.

Table 4.7: Results of post-evaluation tests showing for each best evolved successful group of each evolutionary run ($g*_i$): the average and standard deviation of the fitness over 500 trials in *Env.A* (see columns 2, and 3) and in *Env.B* (see columns 6, and 7); the average and standard deviation of the percentage of timesteps sound was emitted by either robot over 500 trials in *Env.A* (see columns 4, and 5) and in *Env.B* (see columns 8, and 9).

group	*Env.A*				*Env.B*			
	fitness		*signalling (%)*		*fitness*		*signalling (%)*	
	mean	sd	mean	sd	mean	sd	mean	sd
$g*_1$	1.96	0.26	4.64	14.61	1.99	0.07	31.07	3.20
$g*_2$	1.89	0.40	0.00	0.00	1.88	0.24	0.00	0.00
$g*_3$	1.90	0.39	0.00	0.00	1.90	0.35	0.00	0.00
$g*_5$	1.99	0.12	0.00	0.00	1.88	0.26	0.00	0.00
$g*_7$	1.77	0.54	2.97	6.88	1.99	0.01	19.48	2.04
$g*_{10}$	1.94	0.28	3.12	10.11	1.99	0.06	30.29	4.01
$g*_{12}$	1.98	0.45	0.00	0.00	1.91	0.40	0.00	0.00
$g*_{13}$	1.92	0.36	97.10	4.89	1.99	0.04	73.09	4.99
$g*_{15}$	1.93	0.46	5.04	12.64	1.98	0.09	32.47	2.87
$g*_{16}$	1.94	0.22	97.96	3.11	1.96	0.19	75.19	3.91
$g*_{17}$	1.96	0.24	4.12	12.12	1.98	0.13	30.01	4.14
$g*_{18}$	1.97	0.21	3.63	11.59	1.97	0.20	22.23	4.69
$g*_{19}$	1.88	0.43	96.71	3.52	1.99	0.11	77.07	5.23
$g*_{20}$	1.84	0.47	0.00	0.00	1.79	0.51	0.00	0.00

Table 4.8: Further results of post-evaluation tests with the *deaf*, *not-self-sound*, and *not-other-sound* setups for the signalling successful groups $g*_1$, $g*_7$, $g*_{10}$, $g*_{13}$, $g*_{15}$, $g*_{16}$, $g*_{17}$, $g*_{18}$ and $g*_{19}$. For each setup, the table shows the average and standard deviation of the fitness over 500 trials in *Env.A* and in *Env.B*.

group	*deaf*				*not-self-sound*				*not-other-sound*			
	Env.A		*Env.B*		*Env.A*		*Env.B*		*Env.A*		*Env.B*	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
$g*_1$	1.93	0.30	1.02	0.27	1.93	0.33	1.99	0.00	1.94	0.31	0.71	0.34
$g*_7$	1.82	0.46	1.83	0.37	1.79	0.53	1.99	0.00	1.81	0.45	1.83	0.36
$g*_{10}$	1.95	0.22	1.17	0.12	1.94	0.25	1.99	0.00	1.98	0.14	1.17	0.12
$g*_{13}$	1.01	0.79	1.74	0.66	1.93	0.38	1.99	0.09	0.99	0.80	1.75	0.52
$g*_{15}$	1.94	0.32	1.15	0.19	1.95	0.20	1.98	0.10	1.98	0.14	1.15	0.19
$g*_{16}$	0.97	0.55	1.76	0.36	1.93	0.42	1.98	0.12	0.91	0.80	1.75	0.42
$g*_{17}$	1.90	0.33	1.04	0.37	1.96	0.33	1.99	0.00	1.94	0.31	0.88	0.64
$g*_{18}$	1.98	0.17	1.13	0.09	1.96	0.22	1.98	0.17	1.99	0.09	1.13	0.09
$g*_{19}$	0.96	0.72	1.78	0.54	1.91	0.37	1.99	0.07	0.92	0.72	1.77	0.56

robots of these groups with a simple graphical interface, we notice that antiphototaxis is triggered by the absence of non-self produced sound (i.e., input to neuron $N_{10} = 0$) and that the behaviour of the sound output neuron is the opposite of the case of group g_2 (see figure 4.4). Thus, in these groups the external cue that triggers antiphototaxis is the absence rather than the perception of non-self produced sound. Notice that all the above groups make use of process *B3* to discriminate between the two environments, that is, antiphototaxis is triggered by the perception/absence of the tone emitted by their group mates. The behaviour resulting from such a process is characterised by the fact that when a robot takes a decision about the environment and emits a tone (as in e.g., $g{*}_1$) or ceases to emit a tone (as in e.g., $g{*}_{13}$), this tone only triggers antiphototaxis in the other robot of the group. That robot while switching to antiphototaxis will immediately emit a tone itself that will be responsible for triggering antiphototaxis in the robot that first signalled the absence of a *way in* zone.

Concerning the performance of group g_7, it does not undergo any disruption. Thus, the group can be equally successful by employing processes *B1* and *B3*. However, judging from the rather lower rates of success in both environments, we can see that this combination of processes is not perfect. Still, it is an interesting solution that combines solitary and social behaviour, as it can benefit from the social function of sound signalling without exclusively relying on it. In other words, such a behaviour could end up being very robust in cases of, for example, failure to perceive sound. Also, robots controller by such a genotype could categorise the environment even if left alone; should there be social information, they could still exploit it to speed up their decision-making.

The results of these evolutionary runs suggest that evolution only found solutions in which robots perform antiphototaxis in response to a change in the reading of the sound sensor that detects the presence/absence of the non-self produced sound. Signalling groups are favoured over non-signalling groups because the emission of a tone by either robot induces antiphototaxis in the other agent. The beneficial effect of this simple form of communication is linked to the fact that social behaviour induces antiphototaxis in robots which may not have been capable or ready of performing it otherwise. Evolution only produced solutions displaying communication properties—it almost exclusively produced discrimination mechanisms only relying on the perception of non-self-produced sound, thus relying only on the communication resulting from the sound signalling.

Finally, it is interesting to notice that in earlier generations, robots of successful signalling groups seem to pay attention to both self and non-self sound inputs. Later on, evolution filtered this redundancy and came up with a system where agents only pay attention to sound emitted by the other agent. This shows a clear transition from solutions using the process *B4* to solutions using *B3*.

4.7 Discussion

Owing to the properties of our design methodology (i.e., evolutionary robotics), signalling behaviours co-evolved with time-dependent categorisation structures, that is, integration over time. In evolutionary terms, these non-reactive mechanisms might have paved the way for the evolution of signalling. In fact, we can draw some hints from the evolutionary analysis we performed in section 4.6 concerning the evolution of signalling, which suggest that evolution proceeds in an "incremental" way. We observed that signalling was present in the population before successful solutions started to appear, in all the evolutionary runs that produced signalling groups. However, it seemed to have no functional meaning: signals seemed to be produced rather randomly and not with respect to the environmental contingencies. Functional signalling behaviours seem to evolve shortly after evolution produces the first groups able to solve the task without any use of signalling. In other words, communicative solutions seem to be subsequent to non-communicative ones. A possible illustration of this process is that sound production that was previously irrelevant becomes linked to the already evolved mechanisms for environmental discrimination and then, as we have shown in section 4.6, the solutions making use of communication come to outperform those that do not. Another clue in support of these speculations is the comparison of the mechanisms underpinning behaviour in both signalling and non-signalling groups, which was discussed in section 4.4.1. Both solutions rely on an internal neuron integrating sensory information over time; however, for communicative solutions, the sound output behaves similarly.

What we can take from this discussion is that the evolution of signalling seems to be strongly based on already evolved cognitive structures (discrimination capabilities) of the agents (see also Nolfi, 2005). As said above, social solutions to the problem seem to be subsequent to solitary ones. In other words, communication is built upon existing individual behaviour and cognitive structures (categorisation). It should be noted that robots of non-signalling group g_{10} do not ignore sound signals should they be perceived. We tested the reaction of those robots to artificially produced signals and what we found is that the robots stop and spin on the spot once they perceive sound (data not shown). This might indeed suggest that, looking at things from a phylogenetic perspective, the first step towards the evolution of signalling is that robots "learn" to pay attention to their sound inputs. Once such a signalling mechanism evolves, it eventually acquires the social function. In this sense, we could say that the behaviour of group g_{10} looks "precedent" to the behaviour of g_2.

Interestingly, signalling can also evolve should it be deprived of its social function. We have performed an experiment running evolutions with noise only in the initial position of the robots, without any noise in sensors/actuators and ensuring that the robots are initialised in exactly anti-diametrical positions, so to experience exactly the same things during their lifetime (data not shown). What we found is that a small subset of successful strategies relies on the perception of acoustic signals to initiate the action of antiphototaxis. However, as we also have argued in previous sections, the solitary function of sound signalling does not provide any selective advantage to signalling groups.

Instead, it is the social function of signalling (the communication resulting from it) that makes signalling groups more fit than others. In other words, we can attribute the evolution of signalling to its social function, and thus to the effect of emitted signals on other members of the group. This observation justifies the use of the word "signal" in order to describe the emission of sound. In fact, according to Maynard-Smith and Harper (2003), a signal evolves because of its effect on others. A *signal* is defined as *"an act or* *structure that alters the behaviour of another organism, which evolved because the receiver's response has also evolved"*. In contrast, a *cue* is defined as in Hasson (1994): *"a feature of the world, animate or inanimate, that can be used by an animal as a guide to future action"*. Obviously our robots do emit a sound *"as a guide to future action"* (to trigger the action of antiphototaxis), but this is not the reason why signalling behaviours emerged in the first place, even if they also display the latter functionality. Ethologists (see Tinbergen, 1964, for example) considered the existence of cues (or derived activities) as precursors of signals and their subsequent *ritualisation* into signals crucial notions in an effort to explain the evolution of communication. They saw ritualisation as the process of the conversion of a movement or action initially not containing any communicative value into one that does. In our case, this description is absolutely relevant and we could summarise by saying that the individual categorisation seems to be the cue that later on is ritualised into the (acoustic categorisation) signal.

In section 4.6.2 we have seen that signalling groups become more robust as they tend to categorise the environment by initiating antiphototaxis later than non-signalling groups. In other words, we observe that the social context has a bearing and effectively alters the behaviour of the robots with respect to their decision-making. This observation brings to mind examples from zoology and in particular social foraging (mentioned in section 4.1—see Elgar, 1987; Fernandez-Juricic et al., 2005, for examples where the way foraging behaviour changes if the animals are situated in a social context is described). Overall, we can say that the behaviour of the robots is re-shaped (through evolution) as a consequence of the social context in which they are located and the availability at some point in evolution of social information (e.g., categorisation signals).

4.8 Conclusion

In this chapter, we have presented a study on autonomous decision-making which focused on the choice between solitary and social behaviour in a categorisation task. Communication emerged in a system provided with the necessary hardware (i.e., a "mouth" and "ears"), while its use and characteristics were not predetermined by the experimenter, but left to evolution to shape. The evolutionary process autonomously shaped (i) individual categorisation capabilities, (ii) individual signalling capabilities, and (iii) social behaviour.

It turned out that evolution produced signalling behaviour tightly linked to the behavioural repertoire of the agent and that made social behaviour more efficient than solitary behaviour, even though the former was not explicitly rewarded by the fitness function. In fact, as we have discussed in section 4.6, communication serves to increase the robustness and the speed of the categorisation. This study contributes to the understanding of issues concerning the evolution of communication, and, more specifically, the identification of conditions that might facilitate the emergence of communication in populations of embodied agents.

It should definitely be acknowledged that there are elements in our experimental setup that facilitate the evolution of cooperative behaviour, namely (i) the fact that our robot group is composed of genetically identical clones, and (ii) the fact that it is the behaviour of the group that we evaluate after each trial. The above factors leave no room for conflicts of interest or cheating. However, we should stress again that the evolution of signalling and thus cooperation is neither trivial or obvious, in particular because it is not explicitly favoured by the fitness function. The use of a functional and meaningful signalling system is not a question with a binary answer based on chance. For a signalling system to evolve, evolution must produce appropriate signals, appropriate reactions to signals and a reorganisation of the decision-making mechanisms to ensure the robustness of the system. Still, we can certainly ask what might happen if the individuals in the robotic group were not genetically identical, and each robot had a different controller and was evaluated only on the basis of its own performance. Floreano et al. (2007) report on a series of experiments aimed at studying the evolutionary conditions for the emergence of visual communication, and they note that *"under individual selection, the ability to produce visual signals resulted in the evolution of deceptive communication strategies in colonies of unrelated robots and a concomitant decrease in colony performance".* Clearly the aspects of our experimental setup that prevented conflicts of interest and deceptive communication between the robots have had some bearing on our observed results. In future work we intend to withdraw these assumptions and address the broader issue of conditions under which communication may evolve despite the absence of explicitly group-level selection.

In this chapter we have demonstrated the portability of time-dependent decision-making mechanisms that display integration over time properties onto real robots. Even though the controllers were evolved in a simulated world and the simulation did not go as far as implementing possible inter-robot differences (see section 4.5), the system was always successful and we never observed any mistakes in categorising the environment. Our results also show that the use of communication was particularly beneficial in the real world, since the inter-robot differences did in fact severely disrupt the individual decision-making mechanisms of certain agents (see for example *s-bot*$_5$ in the four-robots generalisation test described in section 4.5.2, that would need, if left alone, more than two loops around the light to initiate antiphototaxis). However, it is easy to imagine a case where a robot takes the wrong decision about the state of the world, and initiates antiphototaxis emitting a tone in *Env.A*. That would cause the collapse of the whole system, since all robots would perform the wrong action in that environment, even if their individual discrimination mechanism would have produced a correct categorisation. This event was never observed in reality (see section 4.5), suggesting that the evolved behaviour is also very robust against inter-robot differences. This is an obvious danger that evolution managed to avoid, without the experimenter making this explicit anywhere in the experimental design.

In parallel with studying the effectiveness of the evolved signalling mechanisms when tested on real robots, we found that signalling evolved even in the absence of explicit selected pressure coded in the fitness function. In fact, our analysis revealed a "hidden" benefit for communication. This raises the following issue: should we always equip our robots with "ears" and a "mouth" to make possible the switch from solitary to social behaviour, even in cases when the benefit of communication in such a system is obscure to us, the experimenters? The work presented in this chapter is an example in which communication proves to be beneficial and the evolutionary machinery manages to discover ways to use it by linking it to the rest of the robot behaviour and in the process enhancing the robustness of the system. That is, evolution found an efficient way to use these hardware tools (i.e., "ears" and a "mouth"), and it is our belief that this might also be the case with other more complex tasks as well. We cannot go so far as to claim that incorporating long-range communication devices should be standard, but we can say that evolution seems to discover efficient ways to use them. In other words, by using evolutionary robotics within the context of collective or swarm robotics, such hardware tools for long-range signalling might end up being beneficial for the group's performance on a certain task, even though communication might appear pointless to the experimenter at the time of defining the building blocks of the behaviour.

Finally, the research detailed in this section will serve as the basis for the design of controllers able to produce complex collective behaviour as functional self-assembly in a group of autonomous robots. Indeed, the preliminary experiment presented in chapter 7, section 7.2, is largely inspired by the experimental work that has been presented in this chapter. The major difference will be the response to the individual and collective decision-making; instead of individuals performing antiphototaxis, in the case of functional self-assembly, individual agents should aggregate and assemble in order to overcome an obstacle.

Chapter 5

Experiment II: evolving self-assembly

In this chapter, we report on a series of experiments, performed both in simulation and on real hardware, about the evolution of self-assembly in a group of two robots (the *s-bots*). Our goal is twofold: (i) to show that evolutionary robotics can be applied on real-world tasks of fine sensory-motor coordination such as the physical connection between two robots and (ii) to identify the minimal conditions for self-assembly in autonomous agents and to minimise the assumptions made in previous research works treating this issue.

This chapter will be organised as follows: we start by providing, in section 5.1, a brief background on self-assembly. In section 5.2 we provide a brief review of the state of the art in the area of self-assembling robots and we discuss the limitations of these systems, justifying the methodological choices we have made. In section 5.3 we describe the evolutionary machinery and the experimental scenario used to design neural network controllers. Then, in section 5.4 we show the results of post-evaluation tests on physical robots controlled by the best performing evolved controller and in section 5.5 we shed light on the mechanisms underpinning the behaviour of successful robots. The results presented are discussed in section 5.6 and conclusions are drawn in section 5.7.

5.1 Background

According to Whitesides and Grzybowski (2002), self-assembly is defined as "the autonomous organisation of components into patterns or structures without human intervention". Nature provides many examples of animals forming collective structures by connecting themselves to one another. Individuals of various ant, bee and wasp species self-assemble and manage to build complex structures such as bivouacs, ladders etc. Self-assembly in social insects typically happens in order to accomplish some function (defence, object transport, passage formation etc., see Anderson et al., 2002). In particular, ants of the species *Œcophylla longinoda* can form chains composed of their own bodies which are used to pool leaves together to form a nest, or to bridge a passage between branches in a tree (Hölldobler and Wilson, 1978). Self-assembly is also widely observed at the molecular level (e.g., DNA molecules).

Although ubiquitous in nature, self-assembly remains in general a phenomenon whose operational principles are, both in non-living and living organisms at any scale, not easy to grasp because it is impractical to change many of the parameters that determine the behaviour of the system components (see Whitesides and Grzybowski, 2002). However, self-assembly is particularly appealing to various scientific disciplines. Understanding the mechanisms of self-assembly in the cell may provide further insights to the emergence of life starting from chemical reactions. From an engineering point of view, understanding self-assembly may inspire the design of artificial self-assembling components. The application of such systems can potentially go beyond research in robotics laboratories, space applications being one of the most obvious possibilities (e.g. multi-robot planetary exploration and on orbit self-assembly, see Izzo and Pettazzi, 2007).

5.2 Related work and motivations

In this section we review related work in the area of self-assembling robotic systems which serves as motivation for the research work we present in this chapter.

Several examples of robotic platforms in the literature consist of connecting modules. For a very comprehensive review of self-assembling robotic systems, we direct the reader to the works of Yim et al. (2002a); Groß and Dorigo (2008b); Groß et al. (2006a); Tuci et al. (2006). Following Yim et al. (2002a), it is possible to identify four different categories: chain based, lattice based, mobile and stochastic reconfigurable robots. As this work focuses on the design of autonomous control systems for a mobile self-reconfigurable robotic system (the *swarm-bot*, see chapter 2, section 2.3.1), in the following, we provide a small overview of this category only. We go on to discuss the platform that is used in this study: the *swarm-bot*.

5.2.1 Mobile self-reconfigurable robots

In this class of self-assembling systems, individual units are capable of autonomous sensing and motion; units can act independently in the environment. The first example of a mobile self-reconfigurable robot was the CEBOT (see Fukuda and Nakagawa, 1987; Fukuda and Ueyama, 1994). CEBOT is a heterogeneous system composed of cells with different functions (move, bend, rotate, slide). Even though there are no quantitative results to assess the performance and reliability of this system, Fukuda et al. (1988) have shown how docking can be done between a moving cell and a static object cell with a handcrafted controller. Another robotic system capable of self-assembly is the Super Mechano Colony (Damoto et al., 2001; Hirose, 2001). In this system, autonomous robotic wheels, referred to as child units, can connect to and disconnect from a mother-ship. Yamakita et al. (2003) achieved docking by letting the child unit follow a predefined path. Groß et al. (2006b) recently demonstrated assembly between one and three moving child modules and a static module. Hirose et al. (1996) presented a distributed robot called Gunryu. Each robot is capable of fully autonomous locomotion and the assembled structure proved capable of navigating on rough terrain where a single unit would topple over. However, autonomous self-assembly was not studied as the units were connected from beforehand by means of a passive arm. Self-assembly is also not possible for the Millibot train (see Brown et al., 2002), composed of multiple modules that are linearly linked, since no external sensor has been implemented.

In all the above mobile self-reconfigurable systems, self-assembly is either not achieved at all or is only possible between one unit moving autonomously and a static object/unit. For the sake of consistency, we should also mention two important examples of chain based reconfigurable robots, CONRO and PolyBot. CONRO (Castano et al., 2000) has been used by Rubenstein et al. (2004) to demonstrate autonomous docking between two robots, each composed of two modules and performing motion similar to snake robots. It should be noted, however, that the control was heterogeneous at all levels and the generality of the approach was limited due to orientation and distance constraints. More specifically, the two robots were initially placed at a distance not exceeding 15 cm, with an angular displacement of maximum 45°. Concerning the control, it first achieved alignment through the use of IR sensors and emitters, and then approaching and assembly. Yim et al. (2002b) demonstrated self-assembly with PolyBot: a six-modules arm connected to a spare module on a flat terrain. One end of the arm and the spare module were fixed to the walls of the arena at known positions and the motion of the arm relied on knowledge of the goal position and inverse kinematics. In the first of a three-phases procedure, the arm approached the spare module exploiting the knowledge of the goal position and inverse kinematics. The second phase allowed a further approach and alignment of the arm to the spare module, exploiting IR sensors and emitters. The third phase finally led to the connection to the spare module, which in turn detached from the wall (see Yim et al., 2002b, for more details).

5.2.2 Self-assembly with the *swarm-bot*

The *swarm-bot*, a collective and mobile reconfigurable system (see Mondada et al., 2005; Dorigo, 2005, and chapter 2.3.1, for details), consists of fully autonomous mobile robots called *s-bots*, that can physically connect to each other and to static objects (preys, also called *s-toys*). Groß et al. (2006a) presented experiments improving the state of the art in self-assembling robots concerning the number of robots involved in self-assembly, the generality and reliability of the controllers, and the assembly speed. A significant contribution of this work is in the design of distributed control mechanisms for self-assembly relying only on local perception. In particular, self-assembly was accomplished with a modular approach in which some modules have been evolved and others hand-crafted. The approach was based upon a signalling system which makes use of colours. For example, the decision concerning which robot makes the action of gripping (the *s-bot-gripper*) and which one is gripped (the *s-bot-grippee*) is made through the emission of colour signals, according to which the *s-bots* emitting blue light are assuming the role of *s-bot-gripper* and those emitting red light the role of *s-bot-grippee*. Thus, it is the heterogeneity among the robots with respect to the colour displayed, a priori introduced by the experimenter, that triggers the self-assembly process. That is, a single *s-bot* "born" red among several *s-bots* "born" blue is meant to assume the role of *s-bot-grippee* while the remaining *s-*

bot-grippers are progressively assembling. Once successfully assembled to another *s-bot*, each blue light emitting robot was programmed to turn off the blue LEDs and to turn on the red ones. The switch from blue to red light indicates to the yet non-assembled *s-bots* the fact that a robot has changed role from *s-bot-gripper* to *s-bot-grippee*. This system is therefore based on the presence of a behavioural or morphological heterogeneity. In other words, it requires either the presence of a prey lit up in red or the presence of a robot not sharing the controller of the others, which is forced to be immobile and to signal with a red colour. O'Grady et al. (2005) bypassed this requirement by hand-crafting a decision-making mechanism based on a probabilistic transition between states. More specifically, the allocation of roles (which robot lights up red and triggers the process) depends solely on a stochastic process.

Both research works report a success rate of 100% when trying to achieve assembly between either two robots allowed to move autonomously, or between a robot and a static teammate. Moreover, they report high success rates when the size of the group increases. This high reliability was partly due to a recovery move which was triggered once the robots "realised" that there was some inaccuracy in their effort to connect to each other.

5.2.3 Motivations

The research works presented above have been very successful since they also showed how assembled structures can overcome limitations of the single robots, for instance in transporting a heavy object or in navigating on rough terrain. However, these modularised architectures are based on a set of a priori assumptions concerning the specification of the environmental/behavioural conditions that trigger the self-assembling process. For example, (a) the objects that can be grasped must be red, and those that should be avoided must be blue; (b) the action of grasping is carried out only if all the "grasping requirements" are fulfilled (among others, a combination of conditions concerning the distance and relative orientation between the robots, see Groß et al., 2006a, for details). If the experimenter could always know in advance in what type of world the agents will be located, assumptions such as those concerning the nature of the object to be grasped would not represent a limitation with respect to the domain of action of the robotic system. However, since it is desirable to have agents that can potentially adapt to variable circumstances or conditions that are partially or totally unknown to the experimenter, it follows that the efficiency of autonomous robots should be estimated also with respect to their capacity to cope with "unpredictable" events (e.g., environmental variability, partial hardware failure, etc.). For example, failure to emit or perceive red light for robots guided by the controllers presented above would significantly hinder the accomplishment of the assembly task.

We believe that a sensible step forward in this direction can be made by avoiding to constrain the system to initiate its most salient behaviours (e.g., self-assembly) in response to a priori specified agent's perceptual states. The research work described in this chapter represents a significant step forward in this direction. It illustrates the details of an alternative methodological approach to the design of homogeneous controllers (i.e., where a controller is cloned on each robot of a group) for self-assembly in physical autonomous robots in which no assumptions are made concerning how agents allocate roles. By using dynamical neural networks shaped by artificial evolution, we managed to design mechanisms by which the allocation of the *s-bot-gripper* and the *s-bot-grippee* roles is the result of an autonomous negotiation phase between the *s-bots*, and not predetermined by the experimenter. In other words, the self-assembly process is triggered and regulated by perceptual cues that are brought forth by the agents through their dynamical interactions.

Furthermore, coordination and role allocation in our system is achieved solely through minimal sensors (distance and angle information) and without explicit communication, contrary to the above works where the agents signal their internal states to the rest of the group. Also, due to the nature of the sensory system used, the robots cannot sense the orientation of their group-mates. In this sense, our approach is similar to (and largely inspired from) the one of Quinn (2001); Quinn et al. (2003), where role allocation (leader-follower) or formation movement is achieved solely through infrared sensors. In addition, we also show that the evolved mechanisms are as effective as the modular and hand-coded ones described in (Groß et al., 2006a; O'Grady et al., 2005) when controlling two real *s-bots*.

Finally, we should stress that our design choices are governed by the principles of autonomous decision-making, as they were presented in chapter 2, section 2.7.2.1. The fitness function which is described in section 5.3.4 does not "encode" rules about the way the robots should interact when trying to connect to each other. As mentioned previously, Groß et al. (2006a) and O'Grady et al. (2005) make use of evolved modules to solve the problem of assembly. It is out of the scope of this thesis to argue whether these works use evolutionary robotics as the design methodology; however, they definitely do not share the property of autonomous decision-making with other ER research works. This is because (part of) the interactions among robots are imposed by the experimenter, as we have explained above; the rules of the game are dictated and are not automatically generated.

5.3 Methods

This section is structured as follows: we begin with the description of the task in section 5.3.1, then we describe the simulation model used (see section 5.3.2) and the controller and the evolutionary algorithm chosen (see section 5.3.3) and subsequently we describe the fitness function in section 5.3.4.

5.3.1 The task

The task we consider is conceptually a very simple task. At the beginning of each trial, two *s-bots* are positioned in a boundless arena at a distance randomly generated in the interval $[25\text{ cm}, 30\text{ cm}]$, and with predefined initial orientations α and β (see figure 5.1). Our initialisation is inspired from the initialisation used in (Quinn, 2001). In particular, we defined a set of orientation duplets (α, β) as all the combinations with repetitions from a set:

$$\Theta_n = \left\{ \frac{2\pi}{n} \cdot i \mid i = 0, \ldots, n-1 \right\}, \tag{5.1}$$

where n is the cardinality of the set. In other words, we systematically choose the initial orientation of both *s-bots* drawing from the set Θ_n. The cardinality of the set of all the different duplets—where we consider $(\alpha, \beta) \equiv (\beta, \alpha)$—corresponds to the total number of combinations with repetitions, and can be obtained by the following equation:

$$\frac{(n+k-1)!}{k!(n-1)!}, \tag{5.2}$$

where $k = 2$ indicates that combinations are duplets, and $n = 4$ lets us define the set of possible initial orientations $\Theta_4 = \{0°, 90°, 180°, 270°\}$. From this, we generate 10 different (α, β) duplets, that can be seen in table 5.1.

The goal is to design controllers that can achieve assembly regardless of the initial configuration. More specifically, the robots at the beginning of a trial can be in symmetrical (i.e., $\alpha = \beta$) or asymmetrical configuration (i.e., $\alpha \neq \beta$). Due to our initialisation schema, none of the robots can know at the beginning of the trial if it will become the *s-bot-gripper* or the *s-bot-grippee* at the end of the trial. This is because the setup is neutral towards both α, β and a successful controller has to be successful in all 10 duplets of table 5.1.

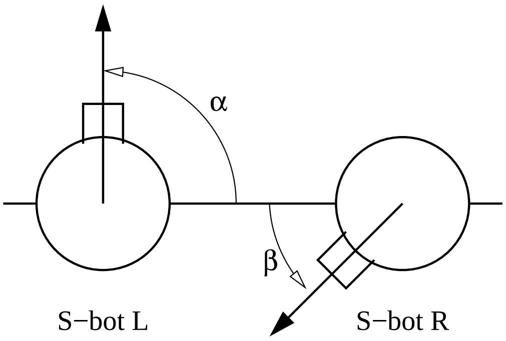


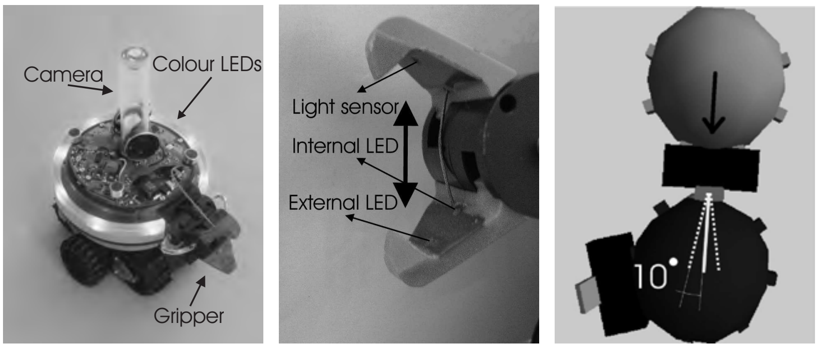
Figure 5.1: This picture shows how the *s-bots*' starting orientations are defined given the orientation duplet (α, β). *S-bot* L and *s-bot* R refer to the robots whose initial orientations in any given trial correspond to the value of α and β, respectively.

Table 5.1: The orientation duplets for the set Θ_4, used during evolution.

	(*s-bot* L, *s-bot* R)		(*s-bot* L, *s-bot* R)
1	$(0°, 180°)$	6	$(90°, 270°)$
2	$(0°, 270°)$	7	$(90°, 0°)$
3	$(180°, 270°)$	8	$(90°, 90°)$
4	$(0°, 90°)$	9	$(180°, 180°)$
5	$(90°, 180°)$	10	$(0°, 0°)$

5.3.2 The simulation model

The controllers are evolved in a simulation environment which models some of the hardware characteristics of the real *s-bots* (see Mondada et al., 2004, and chapter 2.3.1). In this work, to allow robots to perceive each other, we make use of the omni-directional camera mounted on the turret. The image recorded by the camera is filtered in order to return the distance of the closest red, green, or blue blob in each of eight 45° sectors. A sector is referred to as CAM_i, where $i \in \{1, 2, \ldots, 8\}$, denotes the index of the sector. Thus, an *s-bot* to be perceived by the camera must light itself up in one of the three colours using the LEDs mounted on the perimeter of its turret. An *s-bot* can be perceived in at most two adjacent sectors. Notice that the camera can clearly perceive coloured blobs up to a distance of approximately 50 cm, but the precision above approximately 30 cm is rather low. Moreover, the precision with which the distance of coloured blobs is detected varies with respect to the colour of the perceived object. We also make use of the optical barrier which is a hardware component composed of two LEDs and a light sensor mounted on the gripper (see figure 5.2b). By post-processing the readings of the optical barrier we extract information about the status of the gripper and about the presence of an object between the gripper claws. More specifically, the post-processing of the optical barrier readings defines the status of two virtual sensors: a) the *GS* sensor, set to 1 if the optical barrier indicates that there is an object in between the gripper claws, 0 otherwise; b) the *GG* sensor, set to 1 if a robot is currently grasping an object, 0 otherwise. We also make use of the *GA* sensor, which monitors the gripper aperture. The readings of the *GA* sensor range from 1 when the gripper is completely open to 0 when the gripper is completely closed. The *s-bot* actuators are the two wheels and the gripper.[1]



(a) (b) (c)

Figure 5.2: (a) The s-bot. (b) The gripper and sensors of the optical barrier. (c) Depiction of the collision manager. The arrow indicates the direction along which the *s-bot-gripper* should approach the *s-bot-grippee* without incurring into collision penalties.

The simulator used to evolve the required behaviour relies on a specialised 2D dynamics

---

[1] Notice that in this research work we do not make use of the *s-bot*'s degree of freedom for lifting the grasped objects (see chapter 1, section 2.3.1); that is we only allow the rigid gripper to open and close, and not to rotate vertically to the turret (see also figure 5.2b).

engine (see Christensen, 2005). In order to evolve controllers that transfer to real hardware, we overcome the limitations of the simulator by following the approach proposed in Jakobi (1997); motion is simulated with sufficient accuracy, collisions are not. Self-assembly relies on rather delicate physical interactions between robots that are integral to the task (e.g., the closing of the gripper around an object could be interpreted as a collision). Instead of trying to accurately simulate the collisions, we force the controllers to minimise them and not to rely on their outcome. In other words, in case of a collision, the two colliding bodies are repositioned to their previous positions, and the behaviour is penalised by the fitness function if the collision can not be considered the consequence of an accepted grasping manoeuvre.

Concerning the simulation of the gripper, we modelled the two gripper claws as triangles extending from the body of the robot. As the gripper opens, these triangles are pulled in the robot's own body, whereas as it closes they grow out of it. Thus the size of the collision object changes with the aperture of the gripper. In order for a grip to be called successful, we require that there is an object between the claws of the (open) gripper, as close as possible to the interior of the gripper and that the claws close around it. In fact, we require that the object and the gripper socket holding the two claws collide. However, we do not penalise such a collision when the impact angle between the *s-bots* falls within the range [-10°,+10°]. Figure 5.2c shows how this impact angle is calculated and also depicts the simulated robots we use. In this way, we facilitate the evolution of approaching movements directed towards the turret of the robot to be gripped (see figure 5.2c). Robots that rely on such a strategy when attempting to self-assemble in simulation, can also be successful in reality. Other types of strategies based on rotating movements proved prone to failure when tested on real hardware. Having taken care of the collisions involved with gripping, the choice of a simple and fast simulator instead of one using a 3D physics engine significantly speeds up the evolutionary process.

5.3.3 The controller and the evolutionary algorithm

The agent controller is composed of a CTRNN of ten hidden neurons and an arrangement of eleven input neurons and three output neurons (see figure 5.3, Beer and Gallagher (1992) and chapter 3, section 3.1.2 for a more detailed illustration of CTRNNs). Input neurons have no state. At each simulation cycle, their activation values I_i—with $i \in \{1, 2, \ldots, 11\}$—correspond to the sensors' readings. In particular, I_1 corresponds to the reading of the GA sensor, I_2 to the reading of the GG sensor, I_3 to I_{10} correspond to the normalised reading of the eight camera sectors CAM_i, and I_{11} corresponds to the reading of the GS sensor. Hidden neurons are fully connected. Additionally, each hidden neuron receives one incoming synapse from each input neuron. Each output neuron receives one incoming synapse from each hidden neuron. There are no direct connections between input and output neurons. The state of each hidden neuron y_i—with $i \in \{1, 2, \ldots, 10\}$—and of each output neuron o_i—with $i \in \{1, 2, 3\}$—is updated as follows:

$$\tau_i \frac{dy_i}{dt} = -y_i + \sum_{j=1}^{11} \omega_{ji} I_i + \sum_{k=1}^{10} \omega_{ki} Z(y_k + \beta_k); \quad o_i = \sum_{j=1}^{10} \omega_{ji} Z(y_j + \beta_j); \tag{5.3}$$

In these equations, τ_i are the decay constants, ω_{ij} the strength of the synaptic connection from neuron i to neuron j, β the bias terms, and $Z(x) = (1+e^{-x})^{-1}$ is a sigmoid function. τ_i, β, and ω_{ij} are genetically specified networks' parameters. $Z(o_1)$ and $Z(o_2)$ linearly scaled into [-3.2 $\frac{cm}{s}$, 3.2 $\frac{cm}{s}$] are used to set the speed of the left and right motors. $Z(o_3)$ is used to set the gripper aperture in the following way: if $Z(o_3) > 0.75$ the gripper closes; if $Z(o_3) < 0.25$ the gripper opens. Cell potentials are set to 0 when the network is initialised or reset, and circuits are integrated using the forward Euler method with an integration step-size of 0.2.

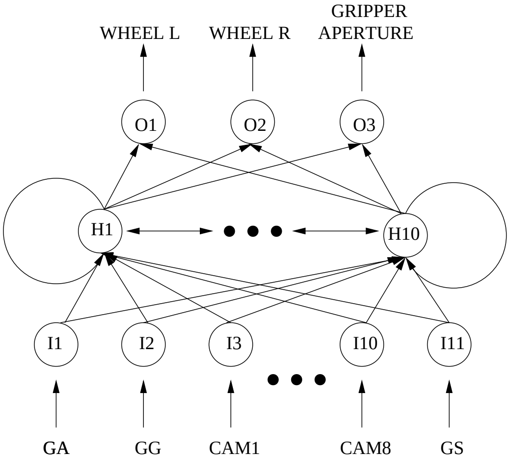


Figure 5.3: Architecture of the neural network that controls the *s-bots*.

Each genotype is a vector comprising 263 real values. Initially, a random population of vectors is generated by initialising each component of each genotype to values randomly chosen from a uniform distribution in the range $[-10, 10]$. The population contains 100 genotypes. Generations following the first one are produced by a combination of selection, mutation, and elitism. For each new generation, the five highest scoring individuals from the previous generation are chosen for breeding. The new generations are produced by making twenty copies of each highest scoring individual with mutations applied only to nineteen of them. Mutation entails that a random Gaussian offset is applied to each real-valued vector component encoded in the genotype, with a probability of 0.25. The mean of the Gaussian is 0, and its standard deviation is 0.1.

5.3.4 The fitness function

During evolution, each group is evaluated 4 times at each of the 10 starting orientation duplets that can be seen in table 5.1 for a total of 40 trials. Each trial (e) differs from the others in the initialisation of the random number generator, which influences the robots initial distance and their orientation by determining the amount of noise added to the orientation duplets (α, β). During a trial, noise affects motors and sensors as well. In particular, uniform noise is added in the range ± 1.25 cm for the distance, and in the range $\pm 1.5°$ for the angle of the coloured blob perceived by the camera. 10% uniform noise is added to the motor outputs $Z(o_i)$. Uniform noise randomly chosen in the range ±5° is also added to the initial orientation of each *s-bot*. Within a trial, the robots life-span is 50 simulated seconds (250 simulation cycles), but a trial is also terminated if the robots incur in 20 collisions. In each trial e, each group is rewarded by the following evaluation function:

$$F_e = A_e \cdot C_e \cdot S_e \tag{5.4}$$

This fitness function seeks to assess the ability of the two robots to get closer to each other and to physically assemble through the gripper. A_e is the aggregation component, computed as follows:

$$A_e = \begin{cases} \frac{1.0}{1.0+atan(\frac{d_{rr}-16}{16})} & \text{if } d_{rr} > 16 \text{ cm;} \\ 1.0 & \text{otherwise;} \end{cases} \tag{5.5}$$

with d_{rr} corresponding to the distance between the two *s-bots* at the end of the trial e;

C_e is the collision component, computed as follows:

$$C_e = \begin{cases} 1.0 & \text{if } n_c = 0; \\ 0.0 & \text{if } n_c > 20; \\ \frac{1.0}{0.5+\sqrt{n_c}} & \text{otherwise;} \end{cases} \tag{5.6}$$

with n_c corresponding to the number of robot-robot collisions recorded during trial e;

S_e is the self-assembly component, computed at the end of a trial ($t = T$ with $T \in (0, 250]$), as follows:

$$S_e = \begin{cases} 100.0 & \text{if } GG(T) = 1\text{, for any robot} \\ 1.0 + \frac{29.0 \sum_{t=0}^{T} K(t)}{T} & \text{otherwise.} \end{cases} \tag{5.7}$$

$K(t)$ is set to 1 for each simulation cycle t in which the sensor *GS* of any *s-bot* is active, otherwise $K(t) = 0$.

Notice that, given the way in which F_e is computed, no assumptions are made concerning which *s-bot* assumes the role of *s-bot-gripper* and which one the role of *s-bot-grippee*. The way in which collisions are modelled in simulation and handled by the fitness function is an element that favours the evolution of assembly strategies in which the *s-bot-gripper* (after assuming its role) moves straight while approaching the *s-bot-grippee* (see section 5.3.2). This has been done to ease transferability to real hardware. The fitness assigned to each genotype is the average fitness achieved in the 40 evaluations.

5.4 Results

As was mentioned at the beginning of this chapter, the goal of this research work is to design through evolutionary computation techniques dynamical neural networks to allow a group of two homogeneous *s-bots* to physically connect to each other. To pursue our objective, we run for 10,000 generations twenty randomly seeded evolutionary simulations. Although several evolutionary runs produced genotypes that obtained the highest fitness score (i.e., $FF = 100$, see section 5.3.4), the ranking based on the evolutionary performances has not been used to select a suitable controller for the experiments with real robots. The reason for this is that during evolution, the best groups may have taken advantage of favourable conditions, determined by the existence of between-generation variation in the starting positions and relative orientation of the robots and other simulation parameters. Thus, the best evolved genotype from generation 5,000 to generation 10,000 of each evolutionary run has been evaluated again on a series of 136,000 trials, obtained by systematically varying the *s-bots*' starting orientations. In particular, we evaluated the evolved genotypes using a wider set of 16 initial orientations Θ_{16}, defined by equation 5.1. This set covers all the possible perceptual configurations for the starting condition of one *s-bot*, which may perceive the other *s-bot* through one or two camera sectors (see figure 5.8 for more details). From this set, equation 5.2 tells us that we can derive 136 different duplets (α, β). Each starting condition (i.e., orientation duplet) was tested in 1,000 trials, each time randomly choosing the robots' distance from a uniform distribution of values in the range $[25 \text{ cm}, 30 \text{ cm}]$. Noise is added to initial orientations, sensors readings and motor outputs as described in section 5.3.4.

The best performing genotype resulting from the set of post-evaluations described above was decoded into an artificial neural network which was then cloned and ported onto two real *s-bots*. In what follows, first we provide the results of post-evaluation tests aimed at evaluating the success rate of the real *s-bots* at the self-assembly task as well as the robustness of the self-assembly strategies in different setups (see section 5.4.1). Subsequently, we illustrate the results of analyses carried out with simulated *s-bots*, aimed at unveiling operational aspects underlying the best evolved self-assembling strategy (see section 5.5).

5.4.1 Post-evaluation tests on real *s-bots*

The *s-bots*' controllers are evaluated four times on each of 36 different orientation duplets (α, β), obtained drawing α and β from Θ_8 (see table 5.2). The cardinality of this set of duplets is given by equation 5.2, with $n = 8$, $k = 2$. In each post-evaluation experiment, successful trials are considered those by which the robots manage to self-assemble, that is, when one robot manages to grasp the other one. Note that, for real *s-bots*, the trial's termination criteria was changed with respect to those employed with the simulated *s-bots*. We set no limit on the maximum duration of a trial, and no limit on the number of collisions allowed. In each trial, we let the *s-bots* interact until physically connected. In a single case we terminated the trial before the robots self-assembled because the *s-bots* moved so far away from each other that they ended up outside the perceptual range of their respective camera. This trial has been terminated after one minute of robot-robot distance higher than 50 cm and it has been considered unsuccessful. As illustrated later in the section, these new criteria allowed us to observe interesting and unexpected behavioural sequences. In fact, the *s-bots* sporadically committed inaccuracies during their self-assembly manoeuvres. Unexpectedly, the robots demonstrated to possess the required capabilities to autonomously recover from these inaccuracies. In what follows, we provide the reader a detailed description of the performance of the real *s-bots* in these post-evaluation trials.[2]

The first two tests with physical robots are referred to as test G25 and test G30. These are tests in which the *s-bots* light themselves up in green and are initialised at a distance from each other of 25 cm and 30 cm, respectively. The *s-bots* proved to be 100% successful in both tests. That is, they managed to self-assemble in all trials. Table 5.3 gives more details about the *s-bots*' performances in these trials. In particular, we notice that the number of successful trials at the first gripping attempt is 28 and 29 trials out of 36 respectively for G25 and G30 (see table 5.3, 2^{nd} column). In a few trials, the *s-bots* managed to assemble after two/three grasping attempts (see table 5.3, 3^{rd} and 7^{th} column). The failed attempts were mostly caused by inaccurate manoeuvres—referred to as inaccuracies of type I_1—, in which a series of maladroit actions by both robots makes impossible for the *s-bot-gripper* to successfully grasp the *s-bot-grippee*'s cylindrical turret. In a few other cases, the group committed a different inaccuracy—referred to as I_2—, in which both robots assume the role of *s-bot-gripper*. In such circumstances, the *s-bots* head towards each other until a collision between their respective grippers occurs. Note that, in both G25 and G30, the *s-bots* always managed to recover from the inaccuracies and end up successful.

As mentioned in section 5.3.2, the *s-bots* have to turn on their coloured LEDs in order to perceive each other through the camera. However, as discussed in section 5.2.1, a significant advantage of our control design approach is that the specific colour displayed

---

[2] Movies of the post-evaluation tests on real *s-bots* can be found at `http://iridia.ulb.ac.be/supp/IridiaSupp2008-002/`.

Table 5.2: The orientation duplets for the set Θ_8, used during the post-evaluations on real hardware. Cells in gray represent the 10 duplets composing the set Θ_4 (used during evolution).

	(*s-bot* L, *s-bot* R)		(*s-bot* L, *s-bot* R)		(*s-bot* L, *s-bot* R)
1	$(0°, 180°)$	13	$(225°, 315°)$	25	$(90°, 225°)$
2	$(0°, 270°)$	14	$(45°, 135°)$	26	$(90°, 315°)$
3	$(180°, 270°)$	15	$(135°, 225°)$	27	$(90°, 45°)$
4	$(0°, 90°)$	16	$(135°, 315°)$	28	$(90°, 135°)$
5	$(90°, 180°)$	17	$(135°, 45°)$	29	$(180°, 225°)$
6	$(90°, 270°)$	18	$(135°, 135°)$	30	$(0°, 45°)$
7	$(90°, 0°)$	19	$(225°, 225°)$	31	$(180°, 45°)$
8	$(90°, 90°)$	20	$(45°, 45°)$	32	$(180°, 135°)$
9	$(180°, 180°)$	21	$(0°, 225°)$	33	$(270°, 225°)$
10	$(0°, 0°)$	22	$(0°, 315°)$	34	$(270°, 315°)$
11	$(45°, 225°)$	23	$(180°, 315°)$	35	$(270°, 45°)$
12	$(45°, 315°)$	24	$(0°, 135°)$	36	$(270°, 135°)$

has no functional role within the neural machinery that brings forth the *s-bots*' actions. In order to empirically demonstrate that the mechanisms underpinning the *s-bots* self-assembling strategies do not depend on the specific colour displayed by the LEDs, we repeated a third and a fourth time the 36 post-evaluation trials, both times by deliberately changing the colour of the *s-bots*' LEDs. The *s-bots* are placed at an initial distance of 30 cm from each other, and they are evaluated with the LEDs displaying blue light—this test is referred to as B30—and with the LEDs displaying red light—this test is referred to as R30.

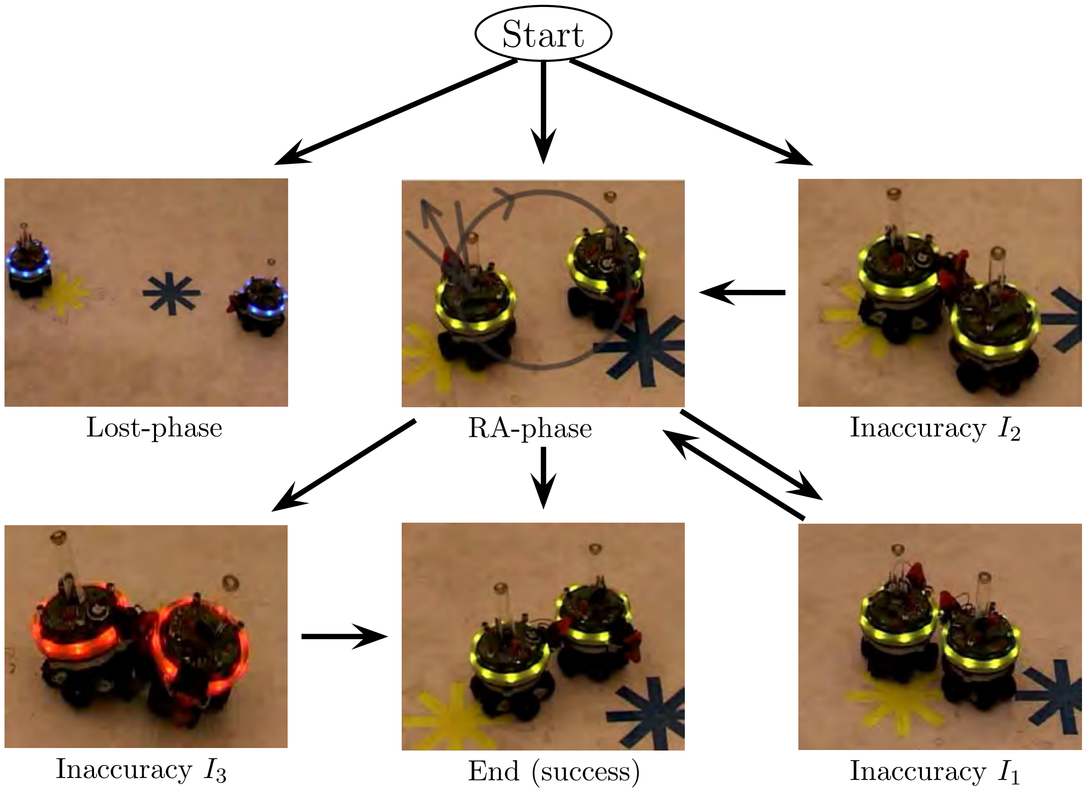
The *s-bots* proved to be very successful both in B30 and R30 (see table 5.3). In the large majority of the trials the *s-bots* managed to self-assemble at the first grasping attempt. In a few trials, two or three grasping manoeuvres were required (see table 5.3, 3^{rd} and 7^{th} column). A new type of inaccuracy emerged in test R30. That is, in three trials, after grasping, the connected structure got slightly elevated at the connection point. We refer to this type of inaccuracy as I_3. Notice also that in a single trial, in test B30, the *s-bots* failed to self-assemble (see table 5.3, last column). In this case, the *s-bots* moved so far away from each other that they ended up outside the perceptual range of their respective camera. This trial in which the *s-bots* spent more than 1 minute without perceiving each other has been terminated, and it was considered unsuccessful.

For each single test (i.e., G25, G30, B30, and R30), the sequences of *s-bots*' actions are

Table 5.3: Results of post-evaluation tests on real *s-bots*. G25 and G30 refer to the tests in which the *s-bots* light themselves up in green and are initialised at a distance from each other of 25 cm and 30 cm, respectively. B30 and R30 refer to the tests in which the *s-bots* light themselves up in blue and red respectively, and are initialised at a distance of 30 cm from each other. Trials in which the physical connection between the *s-bots* requires more than one gripping attempt, due to inaccurate manoeuvres I_i, are still considered successful. I_1 refers to a series of maladroit actions by both robots which makes impossible for the *s-bot-gripper* to successfully grasp the *s-bot-grippee*'s cylindrical turret. I_2 refers to those circumstances in which both robots assume the role of *s-bot-gripper* and collide at the level of their grippers. I_3 refers to those circumstances in which, after grasping, the connected structure gets slightly elevated at the connection point. Failures correspond to trials in which the robots do not manage to return to a distance from each other smaller than their visual field.

Test	Number of successful trials per gripping attempt and types of inaccuracy									*N.°* failures
	1^{st}	2^{nd}				3^{rd}				
	N.°	*N.°*	I_1	I_2	I_3	*N.°*	I_1	I_2	I_3	
G25	28	7	6	1	0	1	2	0	0	0
G30	29	6	3	3	0	1	1	1	0	0
B30	26	5	3	2	0	4	8	0	0	1
R30	20	12	10	0	2	4	7	0	1	0

Table 5.4: Summary of all the possible scenarios observed in real *s-bots* during post-evaluation tests. States refer to particular *s-bots*' spatial configurations that are either representative of phases of a trial (e.g., RA-phase, Lost-phase), specific time in a trial (Start, End), or denote inaccurate manoeuvres that do not hinder the robots from eventually connecting to each other (i.e., I_1, I_2, and I_3, see caption of table 5.3 for further details). The arrows indicate transitions between different states of a trial.



rather different from one trial to the other. However, these different histories of interactions can be succinctly described by a combination of few distinctive phases and transitions between phases which exhaustively "portray" the observed phenomena. Figure 5.4 shows snapshots from a successful trial which represent these phases. The robots leave their respective starting positions (see figure 5.4a) and during the starting phase (see figure 5.4b) they tend to get closer to each other. In the great majority of the trials, the robots move from the starting phase to what we call the role allocation phase (RA-phase, see figure 5.4c). In this phase, each *s-bot* tends to remain on the right side of the other. They slowly move by following a circular trajectory corresponding to an imaginary circle centred in between the *s-bots*. Moreover, each robot rhythmically changes its heading by turning left and right (see table 5.4, middle row, central image). The RA-phase ends once one of the two *s-bots*—that is, the one assuming the role of the *s-bot-gripper*—stops oscillating and heads towards the other *s-bot*—that is, the one assuming the role of the *s-*

bot-grippee—which instead orients itself in order to facilitate the gripping (gripping phase, see figure 5.4d). The *s-bot-gripper* approaches the *s-bot-grippee*'s turret and, as soon as its *GS* sensor is active, it closes its gripper. A successful trial terminates as soon as the two *s-bots* are connected (see figure 5.4e).

5.4.2 The recovery mechanism

As mentioned above, in a few trials the *s-bots* failed to connect at the first gripping attempt by committing what we called inaccuracies I_1 and I_3 (see table 5.4 bottom row, left and right images). These inaccuracies seem to denote problems in the sensory-motor coordination during grasping. Recovering from I_1 can be accomplished by returning to a new RA-phase, in which the *s-bots* negotiate again their respective roles, and eventually self-assemble. Recovering from I_3 is accomplished by a slight backward movement of both *s-bots* which restores a stable gripping configuration. Given that I_3 has been observed only in R30, it seems plausible to attribute the origin of this inaccuracy to the effects of the red light on the perceptual apparatus of the *s-bots*. In particular, it could be that, due to the red light, the *s-bot-gripper* perceives through its camera the *s-bot-grippee* at a farther distance than the actual one. Alternatively, it could be that the red light perturbs the regular functioning of the optical barrier and consequently the readings of the *GS* and *GG* sensors. Both phenomena may induce the *s-bot-gripper* to keep on moving towards the *s-bot-grippee* up to the occurrence of I_3, even though the distance between the robots and the status of the gripper of the *s-bot-gripper* would require a different response.

On the other hand, I_2 seems to be caused by the effects of the *s-bots*' starting positions on their behaviour. In those trials in which I_2 occurs, after a short starting phase, the *s-bots* head towards each other until they collide with their grippers without going through the RA-phase. The way in which the robots perceive each other at starting positions seems to be the reason why they skip the RA-phase. Without a proper RA-phase, the robots fail to autonomously allocate between themselves the roles required by the self-assembly task (i.e., *s-bot-gripper* and *s-bot-grippee*), and consequently they incur in I_2. In order to recover from I_2, the *s-bots* move away from each other and start a new RA-phase in which roles are eventually allocated.

As shown in table 5.3, except for a single trial in test B30 in which the *s-bots* failed to self-assemble, the robots proved capable of recovering from all types of inaccuracies. This is an interesting result because it is evidence of the robustness of our controllers


(a) (b) (c) (d) (e)

Figure 5.4: Snapshots from a successful trial. (a) Initial configuration. (b) Starting phase. (c) Role allocation phase. (d) Gripping phase. (e) Success (grip).

with respect to contingencies never encountered during evolution. Indeed, as mentioned in section 5.3.2, in order to speed up the evolutionary process, the simulation in which controllers have been designed does not handle collisions with sufficient accuracy. In those cases in which, after a collision, the simulated robots had another chance to assemble, the agents were simply re-positioned at a given distance to each other. In spite of this, *s-bots* guided by the best evolved controllers proved capable of engaging in successful recovering manoeuvres which allowed them to eventually assemble.

In short, after an inaccuracy of some type has occurred, the robots have to reconfigure and assume positions that can eventually lead to self-assembly. In most cases, the inaccuracy leads to relative positions (and in general situations) not experienced during evolution, since an inaccuracy of any type (practically always) leads to the termination of the trial due to collisions. Thus, typically, the robots have to escape from these configurations, recreate the conditions that can lead them to successful decision-making, and in essence, repeat the role allocation process. In figure 5.5 we can see a typical example of such a case; after an inaccuracy I_1, the robots initially go away from each other and then renegotiate their roles by repeating the RA-phase.

Still, in few cases, the recovery mechanism practically bypassed the RA-phase since apparently the conditions for decision-making and self-assembly were already there. In figure 5.6 we show snapshots of a trial with backtracking present in the recovery mechanism. We see that after an inaccuracy I_2, one of the robots backtracked and assumed the *s-bot-gripper* role immediately. This is an interesting observation which suggests that during evolution, backtracking was selected for, possibly serving other purposes, like collision avoidance. In case the robots skip the new RA-phase, it takes fewer time for the system to reconfigure and self-assemble.

Figure 5.5: Snapshots from a trial showing the recovery mechanism that involves repetition of the RA-phase.

5.4.3 Comparison of simulated and real robots with respect to time

It is interesting to categorise our results with respect to the time it took for the robots to self-assemble. Figure 5.7 shows the boxplot of trial length for all successful real robot experiments, together with a set of 36 observations extracted from simulation for distances 25 and 30 cm (SimG25 and SimG30, respectively). In simulation, the set of 36 observations

Figure 5.6: Snapshots from a trial showing a recovery mechanism that does not involve a new RA-phase.

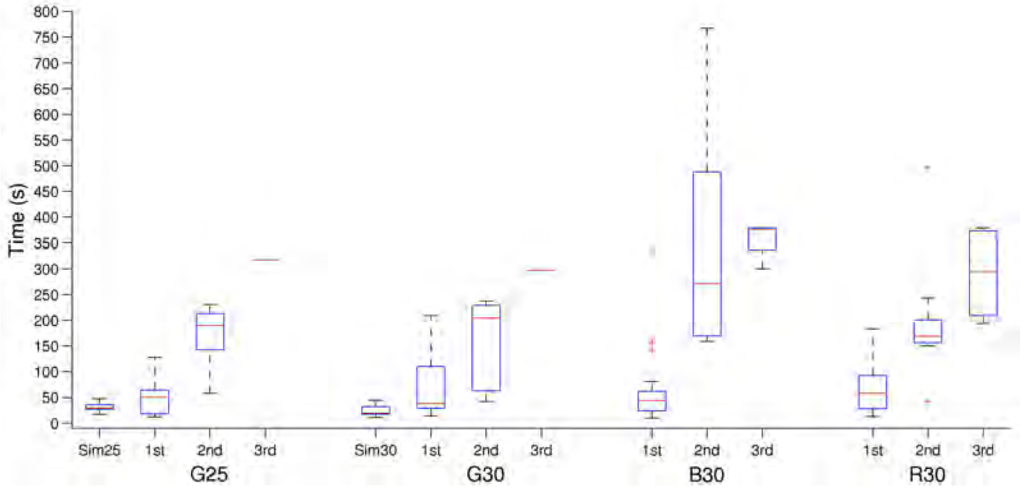


Figure 5.7: Boxplot showing the length of successful trials at the 1^{st}, 2^{nd}, and 3^{rd} gripping attempt per post-evaluation test G25, G30, B30, and R30. The boxplots labelled Sim25 and Sim30 refer to the length of successful simulated trials in which the *s-bots* have been initialised as in the trials corresponding to boxplot labelled G25-1^{st}, and G30-1^{st}, B30-1^{st}, R30-1^{st}, respectively. Boxes represent the inter-quartile range of the data, while the horizontal bars inside the boxes mark the median values. The whiskers extends to the most extreme data points within 1.5 of the inter-quartile range from the box. Crosses denote outliers.

is drawn from Θ_8 and it only contains successful trials (we randomly chose a seed for which all 36 trials end up successful). For every real robot experiment, we group the trials according to the attempt at which assembly was achieved. We will refer to the subset of the experimental results for every experiment as follows: G25-1^{st}, G30-1^{st}, B30-1^{st} and R30-1^{st} will refer to the subset of the corresponding experiment where no inaccuracies were observed, G25-2^{nd}, G30-2^{nd}, B30-2^{nd} and R30-2^{nd} refer to trials where there was one inaccuracy observed and thus the system achieved self-assembly at the second attempt, and G25-3^{rd}, G30-3^{rd}, B30-3^{rd} and R30-3^{rd} refer to trials were assembly was obtained at the third attempt. By comparing the distributions of simulated and real *s-bots* without inaccuracies (Sim25 against G25-1^{st}, and Sim30 against G30-1^{st}, B30-1^{st} and R30-1^{st}), we notice that real *s-bots* tend to take longer time to accomplish their task than simulated agents (Wilcoxon test, 99% confidence interval). It seems that the real-world noise (and possibly other phenomena not accurately modelled in our simulated world, e.g., friction), tends to increase the time required for the *s-bots* to physically assemble. The reader should bear in mind that, in spite of the difference between the length of trials of real and simulated *s-bots*, the robots' controllers proved to be robust enough to accomplish the task with high success rates in both simulation and on real robot hardware. Thus, despite the quantitative differences between the behaviour of simulated and real robots, we consider the experimentation on real hardware successful. Our approach is similar to what described in chapter 4, section 4.5; we again follow Jakobi (1997), who claims that the robots do not have to move identically in simulation and reality in order for the porting to be called successful. In fact, it suffices that real robots carry out the main requirements of the task, that is, in this experiment, to achieve self-assembly. By looking at figure 5.7, we also notice that the coloured light emitted by the LED does not have any clear effect on the length of the successful trials with a single gripping attempt (this is also confirmed by Wilcoxon tests comparing the distributions corresponding to G30-1^{st}, B30-1^{st} and R30-1^{st}—no significant differences were found). As expected, trials that required more than one gripping attempt lasted longer than those in which the *s-bots* managed to assemble at the first attempt (Wilcoxon test, 99% confidence interval).

5.5 An operational description

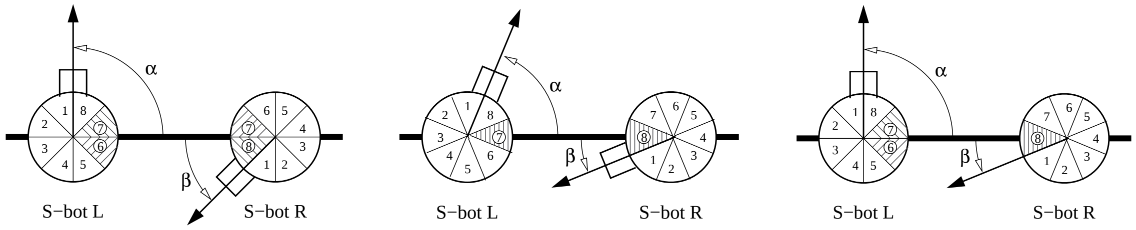
Our research illustrates the details of an alternative methodological approach to the design of controllers for self-assembly in autonomous robots in which no assumptions are made concerning how agents allocate roles in the self-assembly task. The evolved mechanisms are as effective in controlling real robots as those described in (Groß et al., 2006a; O'Grady et al., 2005, see section 5.2.2). Contrary to modular or hand-coded approaches, the evolutionary one proved to be robust with respect to changes of the colour of the light displayed by the LEDs. The controllers described in (Groß et al., 2006a; O'Grady et al., 2005) require to be re-structured by the experimenter in order to cope with the same type of changes.

In view of the results shown in section 5.4.1, we believe that evolved neuro-controllers are a promising approach to the design of mechanisms for autonomous self-assembly. However, it is important to remark that the operational principles of self-assembly used by *s-bots* controlled by our neuro-controllers, are less "transparent" than the modular or hand-coded control described in (Groß et al., 2006a; O'Grady et al., 2005). Further research and experimental analysis are required to unveil the operational principles of the evolved neural controllers. What are the strategies that the *s-bots* use to carry out the self-assembly task? How do they decide who is the *s-bot-gripper*, and who is the *s-bot-*

grippee? Although extremely interesting, providing an answer to this type of questions is not always a simple task. The methodologies at our disposal to uncover the operational mechanisms of evolved neural networks are limited to systems with a small number of neurons, or to cases in which the neural networks control simple agents that only move in a one-dimension world, or by discrete steps (see Beer, 2003a, 2006; Keinan et al., 2006, for examples). Due to the nature of our system, these methods cannot be directly employed to investigate which mechanisms control the process by which two homogeneous *s-bots* differentiate into *s-bot-gripper* and *s-bot-grippee*. In spite of these challenges, we describe in the following the results of an initial series of studies focused on the relationship between the *s-bots*' starting orientations and the role allocation process.

5.5.1 The role of the initialisation

Do the robots' orientations at the beginning of a trial influence the way in which roles (i.e., *s-bot-gripper* versus *s-bot-grippee*) are allocated? We start our analysis by looking at the results of the post-evaluation tests mentioned at the beginning of section 5.4. In particular, we look at those data concerning the behaviour of the *s-bots* controlled by the best performing genotype; that is, the genotype used to build the networks ported on the real robots. Recall that, in these tests, the simulated *s-bots* have been evaluated on a series of 136 starting orientation duplets (α, β) obtained from Θ_{16}. For each orientation duplet, the *s-bots* underwent 1,000 evaluation trials. For each trial, the initial distance was randomly chosen from a uniform distribution of values in the range $[25 \text{ cm}, 30 \text{ cm}]$.



α, $\beta \in \Theta_A$,
$\Theta_A = \{45° \cdot i \mid i = 0, \dots, 7\}$
(a)

α, $\beta \in \Theta_B$,
$\Theta_B = \{45° \cdot i + 22.5°$
$\mid i = 0, \dots, 7\}$
(b)

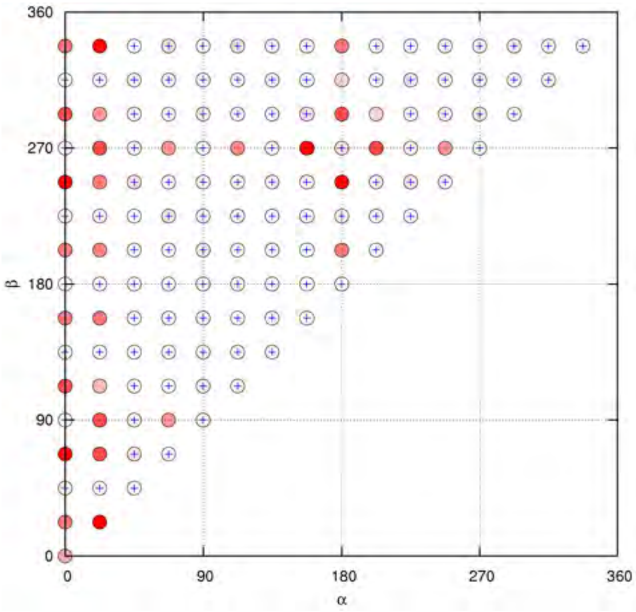
$\alpha \in \Theta_A$, $\beta \in \Theta_B$,
OR
$\alpha \in \Theta_B$, $\beta \in \Theta_A$
(c)

Figure 5.8: Depiction of three different *s-bots*' starting conditions. In each picture, circles represent the *s-bots*, filled arrows indicate the robots' headings and the hollow arrows their orientation. The numbers within the circles refer to the camera sectors CAM_i with $i \in \{1, 2, \dots, 8\}$. Filled sectors are those through which the *s-bots* perceive each other. (a) In trials in which α and β are drawn from the set of starting orientations Θ_A, robots perceive each other in two camera sectors. (b) In trials in which α and β are drawn from the set of starting orientations Θ_B, robots perceive each other in one camera sector. (c) In trials in which $\alpha \in \Theta_i$ and $\beta \in \Theta_j$, with $i \neq j$, robots perceive each other in one and two camera sectors, respectively.

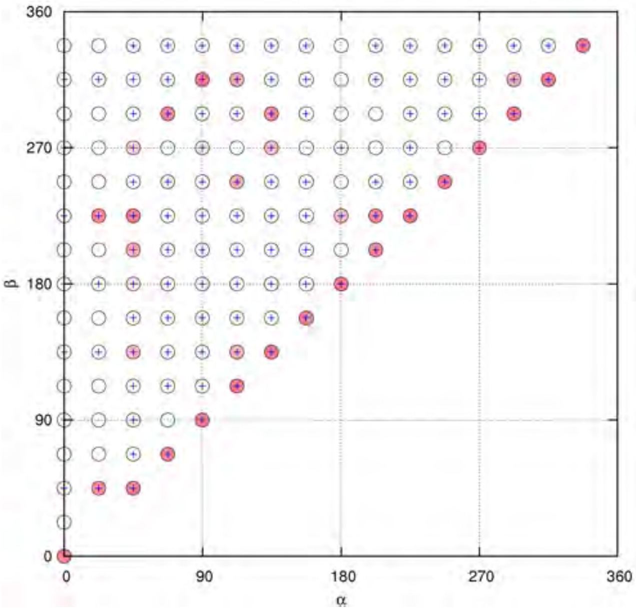
Each duplet (α, β) defines a perceptual scenario at the beginning of a trial characterised by the sector/s through which the robots perceive each other. We defined two subsets of Θ_{16}: Θ_A and Θ_B. These sets encompass those initial orientations that correspond to a perception of the other robot through respectively two camera sectors and one camera sector. As a consequence, each duplet (α, β) identifies one out of three different conditions: $\alpha, \beta \in \Theta_A$ (see figure 5.8a), $\alpha, \beta \in \Theta_B$ (see figure 5.8b), or $\alpha \in \Theta_i, \beta \in \Theta_j$ with $i \neq j$ (see figure 5.8c). Note that, given the distance up to which a coloured blob can be perceived by the *s-bots*' camera (approximately 50 cm), and the dimension of each camera sector (45°), the three categories illustrated in figure 5.8 take into account all the possible perceptual scenarios that the two *s-bots* system can experience at the beginning of a trial, assuming that the robots are at less than 50 cm from each other. The 136 orientation duplets include 16 symmetrical conditions in which $\alpha = \beta$. In symmetrical orientation duplets, the robots share the same perception at the beginning of the trial. That is, they perceive each other through the same sector/s of their corresponding camera. Asymmetrical orientation duplets are those in which $\alpha \neq \beta$. These tests have been repeated twice: once without adding any noise to the robots' orientations (α and β) and once by applying a random offset to the robots' orientations. These offsets are chosen in order not to disrupt the perceptual scenario—i.e., the sector/s through which the robots perceive each other—determined by the corresponding orientation duplet (α, β). The results are qualitatively similar, we therefore in the following discuss only the results of the tests with noise. Recall that, contrary to the real *s-bots*, the simulated robots, due to the way our simulator handles collisions, cannot practically use the recovery manoeuvres observed in the real-world experiments. That is, in these post-evaluation tests, the simulated *s-bots* are scored according to a binary criterion: a trial can be either successful or unsuccessful. Unsuccessful trials are considered those in which the robots did not manage to self-assemble within the time-limit, as well as those that terminated due to the occurrence of collisions that are not considered the result of an accepted grasping manoeuvre (see section 5.3.2 for details).

In figure 5.9a we plot the success rate of the simulated *s-bots*, controlled by the best evolved genotype. In particular, we can see for all the duplets comprising set Θ_{16} the percentage of successful trials. Dark colours refer to low success rates while circles with light colours refer to high success rates. Circles that contain a cross (+) refer to cases where the percentage of success is higher than 85%. The orientation duplets for which the robots self-assembled in 85% or more of the trials amounted to 104 out of a total of 136 duplets. The remaining 32 orientation duplets—1 symmetrical and 31 asymmetrical—contain few cases where the success rate is very low; more specifically, for one duplet ($\alpha = \beta = 22.5°$) we noticed that the success rate was 0%. For this orientation combination the robots always adopted the same role. The three boxes in the boxplot in figure 5.10a correspond to the three initial perceptual scenarios depicted in figure 5.8. We notice that the medians of the three distributions are all above 90%, which indicates that the robots are quite successful in all scenarios. The 2^{nd} and 3^{rd} scenario include some trials where the performance is rather low. However, we cannot conclude that a particular initial perception scenario (e.g., perceiving a robot through one or two camera sectors) has a direct impact on the success rate.

Before proceeding with our analysis of the mechanisms for role allocation, we need



(a)

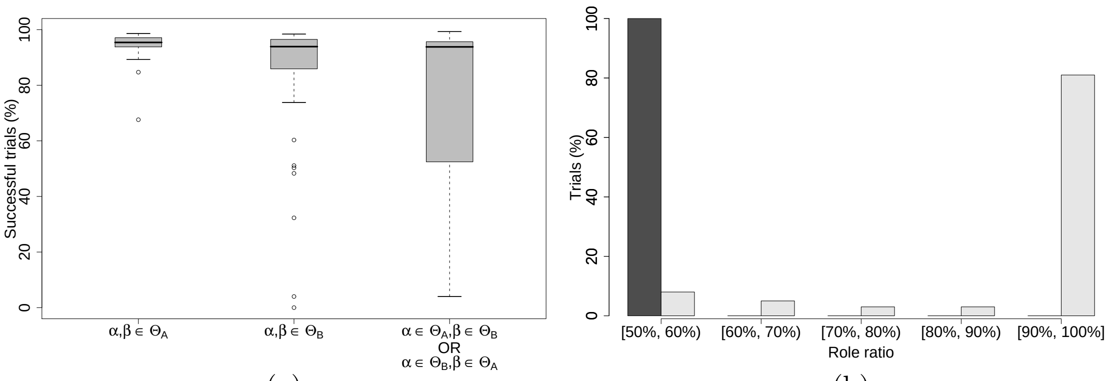


(b)

Figure 5.9: (a) Success rates across all duplets in the set Θ_{16}. The darker the colour of a circle, the lower the percentage of successful trials in the corresponding experiment; (b) The *role ratio* for all duplets of set Θ_{16}. The lighter the colour of a circle, the closer the *role ratio* is to 1 in the corresponding experiment. Circles containing a cross (+) indicate duplets for which the percentage of successful trials is higher than 85%. When $\alpha = \beta = 22.5°$ the success rate is 0% and so there is no data to calculate the *role ratio*.

to clarify the following issues: First, the reader should bear in mind that, in our tests, the *s-bot-gripper* is operationally defined as the robot that successfully grips the other one. This definition has been chosen for being the most reliable in discriminating the roles. Second, given the way in which α and β are varied in our tests, the operational definition of the roles, and the *s-bots* success rate (see figure 5.10a), we can already exclude that the system works by following simple rules by which the role is determined by the initial individual perception. This is because the robots proved to be successful even in symmetrical trials. In other words, having the same initial perception does not hinder the robots from allocating different roles. Therefore, either the system has to be governed by more complex principles based on the combination of α and β, or the initial orientations have no bearing on the role allocation process. In the remainder of this section, we carry out an analysis that helps us further clarify this issue.

For our role allocation analysis we chose only the orientation duplets for which the robots self-assembled in 85% or more of the trials (104 out of 136 duplets). For the remaining 32 orientation duplets for which the robots self-assembled in less than 85% of the trials, we considered the sample size to be too small for a meaningful role allocation analysis. This is mainly due to the operational definition of *s-bot-gripper*. Since there is no *s-bot-gripper* in unsuccessful trials, the lower the *s-bots* success rate the smaller the sample (i.e., the number of trials) in which the relationship between roles and starting perceptual



(a) (b)

Figure 5.10: (a) The plot shows the percentage of successful trials of simulated *s-bots* controlled by the best evolved genotype for each of the three types of starting conditions: (1) α, $\beta \in \Theta_A$, see figure 5.8a; (2) α, $\beta \in \Theta_B$, see figure 5.8b; (3) $\alpha \in \Theta_i$, $\beta \in \Theta_j$, and $i \neq j$, see figure 5.8c. Boxes represent the inter-quartile range of the data, while the horizontal bars 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. Empty circles mark the outliers. (b) The barplot shows the percentage of trials per level of *role ratio*. Black bars refer to orientation duplets (α, β) with $\alpha = \beta$, referred to as symmetrical orientation duplets. Grey bars refer to orientation duplets (α, β) with $\alpha \neq \beta$, referred to as asymmetrical orientation duplets.

scenario can be observed. By looking at the frequency (i.e., how many trials out of 1,000) with which each *s-bot* (i.e., *s-bot* L and *s-bot* R) assumes the role of *s-bot-gripper* for any given value of α and β, our analysis is intended to unveil any relationship between the robots initial orientations and the role they assume during the trial. In particular, we looked at the *role ratio*.

The *role ratio* can be considered a property of each orientation duplet. It indicates how often a given robot (i.e., *s-bot* L or *s-bot* R) assumed the role of *s-bot-gripper* when repeatedly evaluated on a given orientation duplet. In particular, the *role ratio* corresponds to the highest frequency of assuming the *s-bot-gripper* role between the one recorded by *s-bot* L and by *s-bot* R. Thus, the *role ratio* can vary between 50%, when both robots assumed the *s-bot-gripper* role with the same frequency, to 100% when only one robot assumes the *s-bot-gripper* role in all the trials that start with the same perceptual scenario (i.e., evaluation trials repeated for a given orientation duplet). If the *role ratio* is around 50% for both symmetrical and asymmetrical trials, then this is an evidence that the initial individual orientations have no bearing on the role allocation process. If instead, for certain orientation duplets (α, β) the *role ratio* diverges significantly from the 50% value, then we conclude that the system is governed by principles based on the combination of the robots initial individual orientations.

In figure 5.10b, the *role ratio* is divided in five categories, represented on the x-axis, and the orientation duplets are divided in two categories represented by the two types of bars. The black bars refer to the percentage of symmetrical orientation duplets for each category of *role ratio*. The grey bars refer to the percentage of asymmetrical orientation duplets for each category of *role ratio*. By looking at figure 5.10b, we clearly see that while the totality of the orientation duplets corresponding to symmetrical starting positions falls into the category of *role ratio* [50%, 60%), the large majority of the orientation duplets corresponding to asymmetrical starting positions falls into the category of *role ratio* [90%, 100%]. This means that, while in many of the asymmetrical trials the role of *s-bot-gripper* is for the large majority (if not the totality) of the successful trials assumed by the same robot (i.e., either *s-bot* L or *s-bot* R), in all the symmetrical trials both robots assume the role of *s-bot-gripper* with more or less the same frequency. We also found out that in asymmetrical trials with $\alpha = 0°$, it is the *s-bot* L that systematically assumes the role of *s-bot-gripper* (data not shown). For all the other asymmetrical starting conditions with $\alpha \neq 0°$, the role of *s-bot* L depends on the value of β. That is, except for the orientation 0°, for all the other duplets, no associations can be made between an *s-bot*'s initial orientation and its role. In these circumstances (i.e., $\alpha, \beta \neq 0°$), it seems to be the combination of values of α and β which determines whether it is *s-bot* L or *s-bot* R assuming the *s-bot-gripper* role.

For the sake of consistency, in figure 5.9b we plot the *role ratio* for all duplets in Θ_{16}. The darker the colour of a circle, the closer the *role ratio* is to 50%, while very lightly coloured circles refer to *role ratio* close to 1. Circles that contain a cross (+) refer to duplets for which the success rate (plotted in figure 5.9a) is higher than 85%, and that have been taken into account to produce figure 5.10b. Notice that the duplet (α, β)=(22.5°, 22.5°) for which the success rate is 0% is not represented by a circle in the *role ratio* graph. In this figure, we can clearly see that duplets that lie on the diagonal, that is, symmetrical duplets, all have a dark colour denoting a *role ratio* around 50%, while for the great majority of the rest of the duplets the *role ratio* is near 1, as denoted by circles with light colours.

5.5.2 The role of noise

Our analysis revealed that, who is the *s-bot-gripper* and who is the *s-bot-grippee* is the result of an autonomous negotiation phase between the two *s-bots*. The role allocation unfolds in time during the entire duration of a trial. In the case in which the two robots have different initial perceptions, the role that each *s-bot* assumes can (usually) be predicted knowing the combination of α and β. This means that it is this combination which determines the roles. In other words, perceiving the other robot at a specific distance and through a given camera sector or at a certain angle does not inform a robot about the role it assumes during the trial; this role equally depends on the initialisation of the other agent.

However, our results show that in those cases in which the robots start with an identical perception, this symmetry does not hinder the robots from autonomously allocating different roles to successfully accomplish their task. In order to explain this fact, we perform the following test: we systematically reduce the maximum of the random noise applied on sensors and actuators downwards and until no noise is present, and we record the average fitness of the system for 1,000 trials drawn from Θ_8, for three different conditions: (i) asymmetrical ($\alpha \neq \beta$); (ii) symmetrical ($\alpha = \beta$); (iii) symmetrical+1° ($\alpha = \beta + 1^\circ$).

In figure 5.11 we plot the results of this test in a logarithmic scale. We can clearly see that (i) for asymmetrical trials (see continuous line), the noise scaling factor has no effect in performance and the initial asymmetry is what causes the differentiation of the controllers; (ii) for symmetrical+1° (see dotted line), the same as in asymmetrical trials holds, that is, even one degree of difference in the initial perceptions can be enough to produce differentiation; (iii) for symmetrical trials (see dashed line), contrary to the rest of the cases, the noise scaling factor is having a big impact on the success rate, that is, on the amount of trials where differentiation is achieved. In particular, this factor has to be over 10% in order for the performance to reach levels as high as with the other two cases.

Thus, we can say that in the symmetrical case, despite the lack of a priori existing perceptual asymmetries that seem to be guiding the role allocation process, it is the (real-world or injected into the simulation) random noise that injects asymmetries and eventually leads to role allocation. In this particular scenario, stochastic phenomena dominate any causal relationship between environmental structures (i.e., how the robots perceive each other at the beginning of a trial) and the role allocation process.

5.6 Discussion

In a context free of assumptions concerning the nature of the mechanisms underlying the agents behavioural repertoire, our evolutionary robotics model exploits an automatic design process which exploits some features of natural evolution to define the control structures that enable the robots to autonomously self-assemble by assuming complementary roles (i.e., *s-bot-gripper* and *s-bot-grippee*). The results of post-evaluation analyses shown

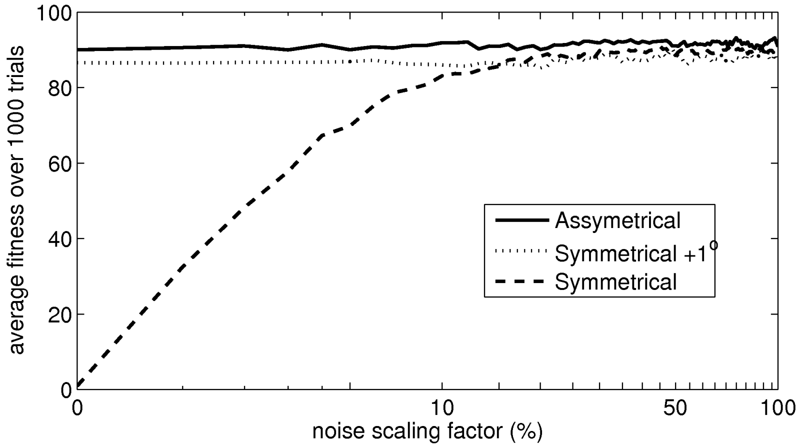


Figure 5.11: Logarithmic plot of the average fitness over 1,000 trials with the noise scaling factor for three different conditions: asymmetrical, symmetrical and symmetrical+1°.

in section 5.5 illustrate that the allocation of roles is the result of an autonomous negotiation phase between the two robots. The outcome of any action an agent chooses depends on the action of the other agent. In other words, none of the two agents can predict its final role from its initial perception.

5.6.1 Minimal conditions for self-assembly

We have shown on real hardware that explicit communication to directly access the "intentions" of the other agents (through explicit signals, as the ones used in Groß et al., 2006a, for example) is not a necessary condition for coordination. Our robots coordinate without direct and explicit communication. Noble (1998) reached a similar conclusion with an evolutionary simulation model involving two simulated animals (controlled by a CTRNN) contesting the possession of a resource. Moreover, in another work, Noble (2000b) studied evolved strategies for communication and action in asymmetrical animal contests and observed that the animats did not use an explicit signalling channel that was available to them, but they rather exchanged information about fighting ability through their movement. Groß and Dorigo (2008a) have also shown that cooperative behaviour can be achieved without explicit means of communication. More specifically, in a cooperative transport task, simulated robots could find effective transport strategies exploiting indirect communication, that is, by interacting with each other indirectly through the object being manipulated. Finally, our results are very similar to the results obtained in (Quinn, 2001; Quinn et al., 2003), where role allocation (leader-follower) and formation movement is achieved solely through infrared sensors and the control structure is once again an evolved dynamic neural network. In particular, the work presented in Quinn (2001) reports on role allocation between two robots for symmetrical and non-symmetrical cases. Whilst the author qualitatively explains how the difference in the initial perceptions influences the role allocation for non-symmetrical cases, an analysis of the evolved behaviour in case of “insufficient differences” is not performed. In the analysis performed in section 5.5, we have explained quantitatively and to some degree qualitatively the effect of the starting configuration on the final outcome of a trial (how roles are allocated); the great majority of non-symmetrical configurations severely bias the role allocation process, while symmetrical configurations are governed by stochastic phenomena that take over causal relationships between initial conditions and the final role allocation.

Achieving coordination in the self-assembly task without access to the intentions or even orientation of the other robot could help as a basis to also reconsider the design of controllers for other group (or swarm) behaviours, for example flocking, where such assumptions are made concerning the cognitive and communication capabilities of the robots (see Kelly and Keating, 1996; Campo et al., 2006; Turgut et al., 2008, for examples). Concerning natural systems where flocking is observed (e.g., insect swarms or fish schools), the principles underlying social interactions and information exchange are still not known (see Couzin et al., 2005; Couzin, 2007). In (Couzin, 2007), the authors present a model of self-organised flocking that does not require the presence of leaders and signals to lead a group to coordinated flocking motion through information exchange. The authors provide the agents with the knowledge of the position and direction vectors of other individuals in a neighbourhood, since it is observed that animals often align their direction of motion with that of nearby neighbours. Our evolutionary robotics model achieves such a coordination with a more minimal setup: an evolved CTRNN can provide a robot with the tools to monitor the activity of another robot, and thus extract useful information (as the direction of motion of the other robot) without assuming this type of information directly available to the robot.

5.6.2 Anti-coordination problems

The kind of scenario described in this chapter requires two agents to assume complementary (opposite) roles in order to autonomously solve a given task. This type of scenarios are also referred to as anti-coordination problems. According to Lewis (1969), in coordination and anti-coordination problems, *“two (or more) agents must choose one of several alternative actions”*. The author continues by stressing that *“the outcome of any action an agent might choose depends on the action of the other agents”*.

Coordination and anti-coordination problems are particularly studied by biologists, either by direct observation of the behaviour of animals in nature or captivity, or by the use of analytical modelling tools, as in Maynard-Smith and Price (1973), where “limited-war” type conflicts (without serious injuries) between conspecifics are studied. Typically, such problems are studied with the use of game theory models (see Maynard-Smith, 1982; Hofbauer and Sigmund, 1998, for examples). The latter models allow biologists to predict the outcome of coordination/anti-coordination problems given the set of behavioural strategies available to the agent and the payoff corresponding to all the possible combinations of actions among the actors. However, such analytical tools seem to be less suitable for testing hypotheses concerning the nature of the underlying mechanisms that underpin the agents’ behaviour. In order to shed light on similar mechanisms, we believe that evolutionary robotics models are suitable modelling tools and complementary to other analytical modelling tools at the disposal of biologists.

In the experiments we have presented, the robots manage to coordinate their actions, "hovering" around the conditions that lead to assuming the *s-bot-gripper* role (oscillatory movement). The way in which the robots solve the anti-coordination problem bears striking similarities to the way in which animals solve conflicts. In fact, the conflict between two strategies as attack and flee (or surrender), approach or avoid, is very common in nature. For example, gulls during fights adopt in turns agonistic (aggressive) postures which are abandoned as the birds turn broadside to the antagonist—a consequence of fear. Eventually, one bird will abandon the offensive and will adopt an appeasement posture or run away (see Tinbergen, 1953, for details). The similarity of this behaviour with the one of our robots is striking: it has even been observed that fighting birds walk parallel or around each other, as our robots circle around each other (see table 5.4, RA-phase). Similar coordination rituals are observed in gulls mating, in the "dance-fighting" observed in the male starling (see Ellis, 1966), in the fighting behaviour but also the mating "zig-zag dance" of the stickleback (Tinbergen, 1952), and in the parallel walks engaged in by red deer stags (Clutton-Brock et al., 1979), which allow for each animal to assess the other's size and strength and to investigate possible asymmetries (Noble, 2000b). Recently, Livnat and Pippenger (2006) showed that *"an optimal decision-making system can involve "selfish" agents that are in conflict with one another, even though the system is designed for a single purpose"*. Similarly, when the robots are in the role allocation phase (see section 5.4.1), their behaviour can be seen as the sum of two conflicting tendencies (assume the *s-bot-gripper* role or not) that co-exist at a dynamic equilibrium and oppose each other at the point of wavering. Finally, parallels can also be drawn between our system and simultaneous hermaphrodites, as snails, slugs and fish species. In those animal species, individuals will have to take single mating decisions ("one-shot" games) that require the coordination with the individual with whom they interact, since assuming the same role will end up costly for both (for a comprehensive review of models that explore the origins and solutions to such conflicts, we direct the reader to Anthes et al., 2006).

Obviously, the way in which our robots solve the self-assembly task is determined by the way in which we set up our evolutionary processes. Understanding how our fitness function constrains the evolution of the *s-bots*' behaviour might help us learn more about the evolved behaviour. With respect to this point, we decided to cast the fitness function into a game-theoretic framework, and in particular to revisit it as the payoff matrix of a simple anti-coordination game (a simple version of the "Hawk-Dove" game,[3] see Maynard-Smith, 1982; Maynard-Smith and Price, 1973). The fitness function selectively rewards the robot group to achieve self-assembly: the robots must coordinate to decide who will grip whom. A failure to take a decision will result in low fitness scores. The same goes if both robots decide to assume the *s-bot-gripper* role; the robots will collide and will be therefore punished by the fitness function. If we make a simplification and we assume that i) if the role allocation is correct, the result is a successful grip and the fitness of the group is maximum (100), ii) a failure in the decision-making (see Inaccuracy I_3 in section 5.4.1)

---

[3] This game which is also known as the game of "Chicken" is based on the principle that the outcome where neither player yields to the other is the worst possible one for both players.

leads to 20 collisions and thus the robots receive a fitness score of 0, and iii) in case of failure to allocate the roles during the length of a trial, robots receive just the aggregation fitness component ($A_e < 1$) (see section 5.3.4 for details), the payoff matrix would be the one in table 5.5.

Table 5.5: The payoff matrix of the game our robots are evolved to play. One robot chooses a strategy from the columns, and the other from the rows. The payoff refers to the fitness score assigned to the group after the end of the trial. A_e is the aggregation component of the fitness function (see equation 5.4).

	s-bot-gripper	*s-bot-grippee*
s-bot-gripper	0	100
s-bot-grippee	100	A_e

Clearly, this problem has only one solution (the action of every agent is optimal according to what the other agent does): that the agents should do the opposite of what the other is doing and thus allocate roles. Notice that this solution is optimal regardless of the selection type (group or individual). In other words, even if we were using heterogeneous robots which were not evaluated collectively but individually, still the optimal solution would be the coordination by role allocation. Given this operational description of the system, we can justify the behaviour exhibited by the robots. More specifically, the circular movement with oscillations described in section 5.4.1 can be seen as the sum of two components: assuming the *s-bot-gripper* role and abandoning it. A premature decision on behalf of one robot to assume the *s-bot-gripper* role might lead to a decision-making error and in the end the robots would end up receiving a fitness score of 0. Thus, a robot has to assume this role while the other agent assumes the *s-bot-grippee* role.

Notice that all we did here was to simplify the description of the fitness function in order to better understand why evolved controllers behave the way they do. Describing it in game theory terms gave us the opportunity to view the experiment from a more high-level point of view and to realise that our experiment was very close to the "Hawk-Dove" game. This clarifies why our results are the way they are and shows that the solution found is not just a random solution in a possible universe; instead, the principles characterising this solution could only be the ones they are.

In an effort to explain the solution to such a simple anti-coordination problem, game theory would go as far as explaining the obvious: agents have to allocate distinct roles. On the contrary, ER models can go beyond the obvious and propose time-dependent mechanisms at the neural level that can produce such role allocation. In this sense, we believe that ER models can complement analytical models of biological studies, as game theory models. However, the use of game theory language can yield limited benefits to our ER experimentation.

5.7 Conclusion

In this chapter, we have presented the results of an evolutionary methodology for the design of control strategies for self-assembling robots. More specifically, to the best of our knowledge, the control method we have proposed for the physical connection of two robots is the only existing in the literature where the role allocation between gripper and grippee is the result of an autonomous negotiation phase between the homogeneous robots; there is no a priori injected behavioural or morphological heterogeneity in the system. Instead, the behavioural heterogeneity emerges from the interaction between the robots. Moreover, the communication requirements of our approach are reduced to the minimum; simple coordination by means of dynamical interactions between the robots—as opposed to explicit communication of internal states—is enough to bring forth differentiation within the group. We believe that reducing the assumptions on necessary conditions for assembly is an important step to obtain more adaptive and more general controllers for autonomous self-assembly.

The results of this work are a “proof-of-concept”: they proved that dynamical neural networks shaped by evolutionary computation techniques directly controlling the robots' actuators can provide physical robots all the required mechanisms to autonomously perform self-assembly. Contrary to the modular or hand-coded controllers described in Groß et al. (2006a); O'Grady et al. (2005), the evolutionary robotics approach did not require the experimenter to make any a priori assumptions concerning the roles of the robots during self-assembly (i.e., either *s-bot-gripper* or *s-bot-grippee*) or about their status (e.g., either capable of moving or required not to move). The evolved mechanisms proved to be robust with respect to changes in the colour of the light displayed by the LEDs. Furthermore, in section 5.4.1 we have presented a system that exhibits recovery capabilities that could not be observed during the evolutionary simulation and that were not coded or foreseen by the experimenter. Such a feature in our case comes for free, while in the case of Groß et al. (2006a) a recovery mechanism had to be designed as a specific behavioural module to be activated every time the robots failed to achieve assembly.

As mentioned in previous sections, our system is not as “transparent” as a hand-coded control system is, as we cannot break its behaviour down to a set of rules or states. Such an endeavour seems to be very challenging and particularly difficult, especially when the network sizes are large and/or the movement of the robots takes place in a continuous and noisy world, such as the real world. However, we would like to stress that we do not consider this step a necessary precondition for the success of research work using evolutionary robotics as a design methodology. Our view is that it is more important to identify those choices that made the implementation and experimentation successful. In other words, we put the stress on better understanding which principles make the evolutionary machinery able to produce efficient rules to guide groups of robots, than on identifying each and every one of these rules.

It should be noted that the robots initialisation is an important parameter for the evolutionary processes. Our choice aimed to evolve a system that can cope with all possible orientation duplets. Altering the proportion of symmetrical and asymmetrical orientation duplets experienced throughout evolution might have an impact on the evolved role allocation strategies. For example, it is possible that presenting evolution with more symmetrical examples can lead to the prevalence of strategies where the role ratio is around 50%. We believe that this is an important and interesting issue to be considered in future work.

Future work will also focus on the scalability of our system. Can the controllers still manage to achieve assembly if there are more than two robots involved? The fitness function rewards two robots connecting to each other but it does not explicitly impose the formation of one single structure: if we put more than three robots in our arena, nothing guarantees the formation of one single *swarm-bot*. We did not perform a full scalability test; however, we did some initial experimentation to see if the evolved controller tested on the real robots can produce self-assembly with three robots. The results are encouraging and we show the outcome of an example trial in figure 5.12. We see that one of the robots "chooses" one neighbour over the other to grip, and after two robots assemble, the third robot also becomes member of the structure; interestingly, the robots did not go through the role allocation phase. We consider these hints that the system could potentially scale, and this can lead from the study of self-assembly among *s-bots* to the study of self-assembly among *swarm-bots*; however we wish to address this issue in future work taking into account important issues that have been disregarded in the current work. More specifically, the connected structure must have the ability to move coordinately: it should be able to perform coordinated motion (see Baldassarre et al., 2007), which means it should be equipped with more sensors and actuators (traction sensor and rotating turret for the case of the *s-bot*), in order to actively participate in the assembly process. For example, it could interact with other assembled structures or individual robots by either receiving connections from them or grasping them. Also, if in any case scalability is a desired property of our controllers, then it might be more useful to run new evolutionary runs with more than two robots participating in the trials. This is because, the controllers we evolved may be "optimised" for or targeted only to a two-robot case. This is a common problem to ER experimentation; for example, the controllers evolved in (Quinn et al., 2003) prove to be non-scalable; that is, they cannot successfully control a group of robots whose cardinality is larger than the one with which they were evolved (see also Vicentini and Tuci, 2007). However, we might expect the system to be able to cope with this challenge to a certain extent due to its tolerance to inaccuracies and the recovery mechanism.

Figure 5.12: Snapshots from an example trial with 3 robots.

Finally, it should be mentioned that the research detailed in this section will be integral to the study of functional self-assembly, that will be tackled in future work. More in detail, we have managed to find a simulated environment that can sufficiently model the relevant and important aspects involved with the fine-grained sensory-motor coordination required in order to achieve assembly on real robots. Thus, we will reuse this environment in order to study cases when the assembly should not be a priori demanded, but instead, when it should be a consequence of the environmental contingencies. In chapter 7, section 7.2 we present initial experimentation aimed to study the switch to collective behaviour in a functional self-assembly scenario, that is, a scenario where the assembly is functional to the achievement of some goal. That experiment will be an adaptation of the experiment presented in chapter 4.

Chapter 6

Experiment III: evolving communication in morphologically heterogeneous robots

This chapter is dedicated to the presentation of a research work aimed to investigate a particular scenario that requires a group of robots to use communication in order to perform a collective navigation task. In particular, our objective is to prove that evolutionary robotics methods can be successfully applied to the design of homogeneous controllers for morphologically different groups of robots. Using ER, we want to obtain autonomous time-based decision-making and communication in a heterogeneous group of robots, where the role of each robot in the group should not be a priori defined by the experimenter, but should emerge from its dynamical interaction with the rest of its teammates.

The organisation of this chapter is as follows: we start by providing the background, in section 6.1, and the related work, in section 6.2, of the research work detailed in this chapter. We continue with the presentation of the experimental setup in section 6.3, containing the description of the task, the simulation model, the controller and the evolutionary algorithm, and the fitness function. Results of our evolutionary simulation will be presented in section 6.4, and the chapter closes with discussion and conclusions in sections 6.5 and 6.6, respectively.

6.1 Background

In ER, the homogeneous approach (i.e., robots of a group that share the same controller cloned on each agent) is extensively used to deal with morphologically identical robots. This approach is preferred to alternative solutions because it facilitates the design process. For example, the evaluation of the collective behaviour of various homogeneous groups can be directly used to quantitatively estimate the effectiveness of their control structures and subsequently to compare them (see for example the experiments detailed in chapters 4 and 5).

On the contrary, if the robots of a group do not share the same controller, it becomes less intuitive to define the criteria to estimate, from the observation of the collective behaviour, the effectiveness of each control structure within a group and to compare different controllers associated with different groups (see Quinn, 2001). Moreover, the homogeneity of control structures does not preclude the emergence of behavioural specialisations. For example, the work of Quinn et al. (2003) shows that leader/follower specialisation can be obtained in a homogeneous group of robots by using dynamical neural network controllers. Also, the experiment presented in chapter 5 demonstrated that role allocation can be achieved with the use of a CTRNN cloned on two agents. In that chapter, we showed that the same controller, when downloaded on robots with identical morphology, can exploit their history of interactions, noise in the sensors/actuators and potential differences among the robots hardware in order to cause their specialisation or allocation of distinct roles. In this case, we will explore the same approach, that is, the same dynamical neural network ported on a group of real robots, but in order to study specialisation in a group of heterogeneous robots.

To the best of our knowledge, the homogeneous approach has never been employed in the context of morphologically different robots. In this latter context, the neuro-controller may be required to deal with specialisation or role allocation among morphologically similar robots, as well as to account for the morphological differences which characterise the sensory-motor apparatus of the agents. In this work, the process by which a single controller adapts to morphologically different robots is referred to as “dynamic speciation”. The fact that the dynamic neuro-controller can take time into account to produce the robot’s actions (robots can display time-based decision-making) plays an important and central role in the differentiation of the controllers according to the morphological particularities of each type of robot. Given the nature of the adaptive task described in this chapter, we decided to use it as a test-bed to explore the potentiality of the homogeneous approach to design controllers for morphologically different robots.

In accordance to the principles of the design of experiments laid down in chapter 2, section 2.7.2.1, our experimental setup will remain unbiased towards the nature/type of communication and its relation to the rest of the behavioural repertoire of the robots. Note also that designing a single controller for robots with different sensory systems is not a trivial engineering task: this constitutes one more reason to resort to an automated process.

6.2 Related work

Probably the best known control architecture for heterogeneous robots is ALLIANCE (Parker, 1998). ALLIANCE is a distributed, behaviour-based architecture that allows for the control of teams of heterogeneous robots performing tasks consisting of loosely coupled subtasks that could have ordering dependencies. ALLIANCE allows teams of robots to adaptively choose between available actions, according to the state of the environment, the actions of other robots but also each robot’s internal states. A drawback of ALLIANCE is the fixed structure among basic behaviours, which requires the designer to a priori define the dependencies between behaviours.

Recently, Tang and Parker (2005) presented ASyMTRe, a methodology for automatically synthesising task solutions for heterogeneous multi-robot teams. In contrast to prior approaches where the designer was required to specify the role allocation in a given task, ASyMTRe automates how solutions are generated. This is done by the automatic synthesis of multi-robot behaviours for accomplishing the team objectives, by putting together basic behavioural schemas. It is important to stress that the method proposed provides a mechanism for sharing sensory information among robots, so that more capable robots can assist less capable robots in accomplishing their objectives.

The above works use behaviour-based robotics or basic pre-coded behavioural schemas as part of the design methodology; instead, in the work presented in this chapter, we will use evolutionary robotics to design the individual controllers of the heterogeneous robots.

Reference should also be made to (Werner and Dyer, 1992), where the authors evolved neural networks to control male and female agents, due to the conceptual similarities of that work and ours. The task was to co-evolve a population of embodied simulated males and females who live in a grid world and can agree on the interpretation of signals emitted only by the females. Females are supposed to guide the behaviour of perceived males, while the latter are blind and can only perceive the signals emitted by the females. Indeed, communication evolves and the authors also notice the evolution of different, competing dialects. In this work, there are agents (males) deprived of some sensory capabilities (they are blind), thus communication concerns their guidance by the females with full perception capabilities. In contrast to (Werner and Dyer, 1992), in our experiments both types of robots lack full sensory capacities and they have to complement each other in order to fulfil the requirements of the task they will be confronted with.

There are a number of works carried out in the recent past in which agents asked to solve rather simple tasks that require cooperation and coordination develop simple forms of ritualised social interactions and/or signalling capabilities. We have previously reviewed this literature in chapter 3, section 3.2, and we direct the reader to that chapter in order to get a full flavour of the field. However, we repeat here the description of a couple of research works that serve as direct inspiration for the experiment to be presented in the following of this chapter.

In the work described in (Quinn, 2001; Quinn et al., 2003), a team of robots is required to move in an arbitrarily chosen direction by remaining at a distance from each other smaller than the range of their infrared sensors. This work is particularly important because it shows that it is possible to design, through artificial evolution, neural mechanisms which, by simply using the infrared sensors readings, allow a group of homogeneous robots to engage themselves in social interactions which result in the emergence of roles such as leader/follower. The authors also describe the evolution of communication among the robots by showing that behaviour for social coordination first evolves in a non-communicative context, and only subsequently acquires its adaptive function.

In (Di Paolo, 2000), the author describes an experiment in which two autonomous agents, equipped with sound sensors and effectors, have to remain close to each other as long as possible. An operational description of the best evolved solutions reveals that the sound signalling system is used by the agents both for self-stimulation and for social interaction. This evidence seems to go against a shared perspective in biology/psychology which tends to distinguish the behaviour of natural organisms in socially and non-socially relevant. The author uses the counter-intuitive result of his analysis to point out the importance of grounding the functional description of the behaviour of natural organisms into *"what we know about the operation (at different level) of the system concerned"* (Di Paolo, 2000).

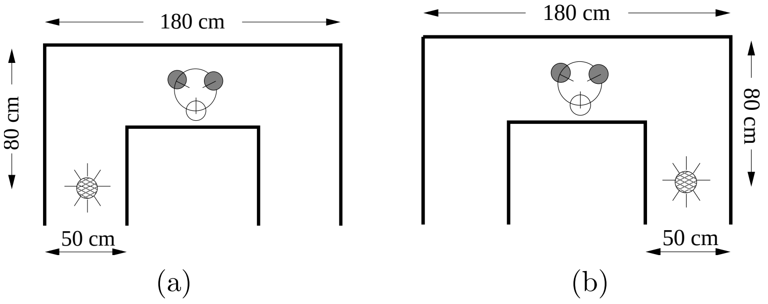
The experiment presented in this chapter is inspired by (Di Paolo, 2000) when it comes to modelling the sound signalling system of our robots (see section 6.3.2). Furthermore, we draw inspiration from Quinn et al. (2003), in order to design the fitness function described in section 6.3.4. Other aspects of our work such as the nature of the cooperative scenario used to investigate issues concerning the evolution of acoustic communication (i.e., social interactions in morphologically heterogeneous robots) and other methodological choices are original and innovative. In section 6.5, we will come back to the similarities and differences between our work and those described in this section, and we will revisit the novel and interesting parts of this research work.

6.3 Methods

This section is structured in the following way: first we describe the task we consider in section 6.3.1. We go on to present the simulation model used (see section 6.3.2) and the controller and the evolutionary algorithm chosen (see section 6.3.3), and we close by describing the fitness function in section 6.3.4.

6.3.1 Description of the task

We consider the following experiment: three simulated robots are required to navigate towards a light source, while remaining close to each other. The robots are placed in an arena, as shown in figure 6.1. The arena is composed of walls and a light that is always turned on. The light can be situated at the bottom left corridor (*Env. L*) or at the bottom right corridor (*Env. R*). The robots are initialised with their centre anywhere on an imaginary circle of radius 12 cm centred in the middle of the top corridor, at a



(a) (b)

Figure 6.1: (a) *Env. L*; (b) *Env. R*. In both pictures, the thick lines represent the arena walls; the two small filled circles represent robots R_{IR}, the small white circle represents robot R_{AL}; the light is represented by the filled circles at the bottom left/right.

minimum distance of 3 cm from each other. Their initial orientation is always pointing towards the centroid of the group. The goal of the robots is (i) to navigate towards the light whose position changes according to the type of environment they are situated in, and (ii) to avoid collisions.

The peculiarity of the task lies in the fact that the robots are equipped with different sets of sensors. In particular, two robots are equipped with infrared and sound sensors but they have no ambient light sensors. These robots are referred to as R_{IR} (see figure 6.2a). The other robot is equipped with ambient light and sound sensors but it has no infrared sensors. We refer to this robot as R_{AL} (see figure 6.2b). Robots R_{IR} can perceive the walls and other agents through infrared sensors, while the robot R_{AL} can perceive the light. Therefore, given the nature of the task, the robots are forced to cooperate in order to accomplish their goal. In fact, it would be very hard for each of them to solve the task solely based on their own perception of the world. R_{AL} can hardly avoid collisions; R_{IR} can hardly find the light source. Thus, the task requires cooperation and coordination of actions between the different types of robots.

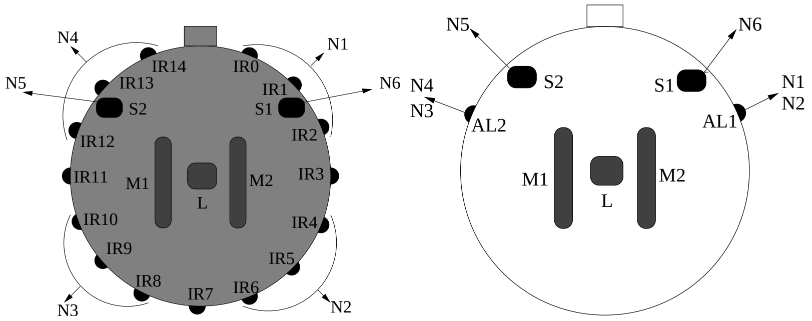
Although the robots differ with respect to their sensory capabilities, they are homogeneous with respect to their controllers. That is, the same controller, synthesised by artificial evolution, is cloned on each member of the group. Both types of robots are equipped with a sound signalling system (more details in section 6.3.2). However, contrary to other studies (see Nolfi, 2005; Marocco and Nolfi, 2006b; Baldassarre et al., 2003), we do not assume that the agents are capable of distinguishing their own sound from that of the other agents. The sound broadcast into the environment is perceived by the agent through omnidirectional microphones. Therefore, acoustic signalling is subject to problems such as the distinction between own sound from that of others and the mutual interference due to lack of turn-taking (see Di Paolo, 2000).

Notice that the reason why we chose the group to be composed of two R_{IR} and one R_{AL} robot is that this intuitively seems to be the smallest group capable of spatially arranging itself adaptively in order to successfully navigate in the environment. Preliminary studies have shown that with groups of two robots, evolution tends to favour solutions in which, during navigation, R_{IR} remains in front of R_{AL}. This type of group has troubles in making the left and the right turn. As we will show in the next sections of this document, a three robot group in which R_{AL} tends to remain behind the two R_{IR} fellows, employs safer and more robust navigation strategies, that allow the robots to successfully make both turns.

6.3.2 The simulation model

The controllers are evolved in a two-dimensional simulation environment which models the kinematics of simple geometries and the functional properties of three types of sensors: infrared, ambient light, and sound sensors (see Vicentini and Tuci, 2006, for a detailed description of the simulator). As illustrated in figure 6.2a and 6.2b, our robots are modelled as circular objects of 5.8 cm of radius. Differential Drive Kinematics equations, as presented in (Dudek and Jenkin, 2000), are used to update the position of the robots within the environment.

The morphological structure and sensory apparatus of our robots model some of the



(a) (b)

Figure 6.2: (a) The robots R_{IR}; (b) The robots R_{AL};

characteristics of the *s-bots* (see Mondada et al., 2004, and chapter 2.3.1, for details). Each robot R_{IR} has 15 infrared sensors (IR_i) placed on the perimeter of its circular body (see figure 6.2a). Robot R_{AL} has two ambient light sensors (AL_1) and (AL_2) positioned at $\pm 67.5°$ with respect to its facing direction (see figure 6.2b). The signal of both infrared sensors and ambient light sensors is a function of the distance between the robot and the obstacle. Both R_{IR} and R_{AL} robots are equipped with a loudspeaker (L) that is situated in the centre of the body of each robot, and with two omnidirectional microphones (S_1 and S_2), placed at $\pm 45°$ with respect to the robot's heading. Sound is modelled as an

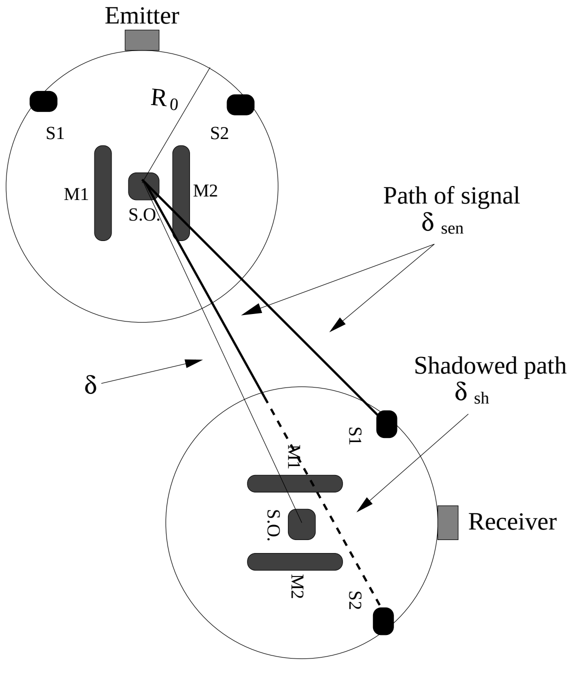


Figure 6.3: This picture has been adapted from (Di Paolo, 2000). It shows the working principles of the shadowing mechanism.

instantaneous, additive field of single frequency with time-varying intensity ($\eta_i \in [0.0, 1.0]$) which decreases with the square of the distance from the source, as previously modelled in (Di Paolo, 2000). Robots can perceive signals emitted by themselves and by other agents. The modelling of the perception of sound is inspired by what described in (Di Paolo, 2000). There is no attenuation of intensity for self-produced signal. The perception of sound emitted by others is affected by a "shadowing" mechanism which is modelled as a linear attenuation without refraction, proportional to the distance (δ_{sh}) travelled by the signal within the body of the receiver (see Di Paolo, 2000, for details). This distance is computed as follows:

$$\delta_{sh} = \delta_{sen}(1 - A), \qquad 0 \leq A < 1, \qquad A = \frac{\delta^2 - R^2}{{\delta_{sen}}^2} \tag{6.1}$$

where δ_{sen} is the distance between the sound source and the sensor, δ is the distance between the sound source and the centre of the body of the receiver, and R is the robot's radius (see also figure 6.3). The *self* component of the sound signal is simply equal to η_i. In order to calculate the *non-self* component, first we scale the intensity of sound emitted by the sender (η_j) by applying the inverse square law with respect to the distance between the sound source and the microphones of the receiver. Subsequently, we multiply the scaled intensity with an attenuation factor ψ which ranges linearly from 1 when $\delta_{sh} = 0$ to 0.1 when $\delta_{sh} = 2R$. To summarise, the reading $\hat{S}_{is}$ of each sound sensor s of robot i is computed as follows:

$$\hat{S}_{is} = \text{self} + \text{non-self}; \qquad \begin{aligned} &\text{self} = \eta_i \\ &\text{non-self} = \sum_{\substack{j \in \{1,2,3\} \\ j \neq i}} \eta_j \frac{R^2}{\delta_{sen}^2} \psi \end{aligned} \tag{6.2}$$

The auditory receptive field of each microphone is bounded within the interval $[0.0, 1.0]$. Therefore, the sound sensor can be saturated by the self emitted sound in case a robot emits at its highest intensity ($\eta_i = 1.0$).

10% uniform noise is added to all sensor readings, the motor outputs and the position of the robot.

It should be reminded that sound is modelled as an instantaneous, additive field of single frequency with time-varying intensity which decreases with the square of the distance from the source, as previously modelled in (Di Paolo, 2000)—a research work also constrained in simulation. Notice that even if this model seems biologically plausible, and even if we did not assume directional microphones or built-in or hard-wired sophisticated discrimination mechanisms, its applicability on the available robotic hardware (i.e., the *s-bot*) seems to be not feasible. This is mainly due to the fact that perception of sound is particularly difficult to simulate accurately, due to noise in the environment and reflections of the sound waves on the walls of the experimental room. In the case of the experiment detailed in chapter 4, sound was modelled as a beep at a single frequency, which can easily be implemented on the *s-bot*. However, we believe that the sound model employed is

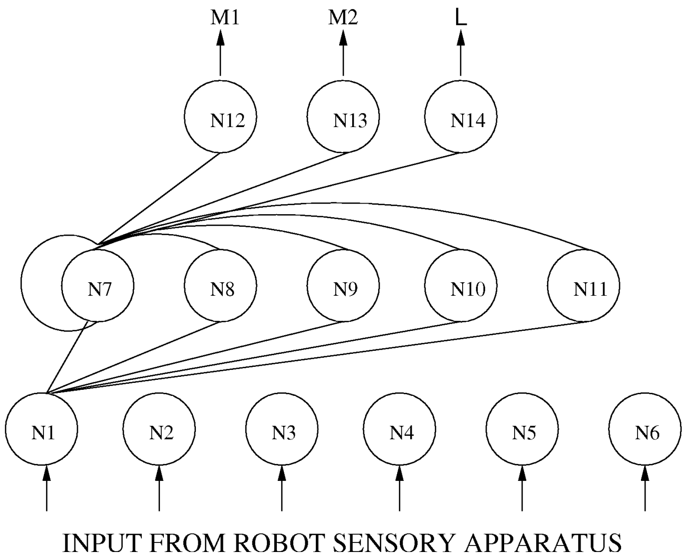


Figure 6.4: The network architecture. Only the efferent connections for one neuron of each layer are drawn. The input layer of R_{IR} takes readings as follows: neuron N_1 takes as input $\frac{IR_0+IR_1+IR_2}{3}$, $N_2 \leftarrow \frac{IR_4+IR_5+IR_6}{3}$, $N_3 \leftarrow \frac{IR_8+IR_9+IR_{10}}{3}$, $N_4 \leftarrow \frac{IR_{12}+IR_{13}+IR_{14}}{3}$, N_5 from sound sensor S_2, and N_6 from sound sensor S_1. The input layer of R_{AL} takes readings as follows: N_1 and N_2 take input from ambient light sensors AL_1, N_3 and N_4 take input from AL_2, N_5 from S_2, and N_6 from S_1. M_1 and M_2 are respectively the left and right motor. L is the loudspeaker.

the simplest possible that can provide the heterogeneous robots under consideration the communication means to accomplish their task.

6.3.3 The controller and the evolutionary algorithm

The agent controller is composed of a network of five inter-neurons and an arrangement of six sensory neurons and three output neurons (see figure 6.4). The sensory neurons receive input from the agent sensory apparatus. Thus, for robots R_{IR}, the network receives the readings from the infrared and sound sensors. For robots R_{AL}, the network receives the readings from the ambient-light and sound sensors. The inter-neuron network (from N_7 to N_{11}) is fully connected. Additionally, each inter-neuron receives one incoming synapse from each sensory neuron. Each output neuron (from N_{12} to N_{14}) receives one incoming synapse from each inter-neuron. There are no direct connections between sensory and output neurons. The network neurons are governed by the following state equation:

$$\frac{dy_i}{dt} = \begin{cases} \frac{1}{\tau_i}(-y_i + gI_i) & i \in \{1,2,\ldots,6\} \\ \frac{1}{\tau_i}\left(-y_i + \sum\limits_{j=h}^{k} \omega_{ji}\sigma(y_j + \beta_j)\right) & i \in \{7,8,\ldots,14\};\ \sigma(x) = \frac{1}{1+e^{-x}} \end{cases} \tag{6.3}$$

where, using terms derived from an analogy with real neurons, y_i represents the cell potential, τ_i the decay constant, g is a gain factor, I_i the intensity of the sensory perturbation on sensory neuron i, ω_{ji} the strength of the synaptic connection from neuron j to neuron i, β_j the bias term, $\sigma(y_j + \beta_j)$ the firing rate. For each i the indexes h and k are set by taking into account the network architecture. The cell potentials y_i of the 12^{th} and 13^{th} neuron, mapped into [0.0,1.0] by a sigmoid function σ and then linearly scaled into $[-6.5, 6.5]$, set the robot motors output. The cell potential y_i of the 14^{th} neuron, mapped into $[0.0, 1.0]$ by a sigmoid function σ, is used by the robot r to control the intensity of the sound emitted η_r. The following parameters are genetically encoded: (i) the strength of synaptic connections ω_{ji}; (ii) the decay constant τ_i of the inter-neurons and of neuron N_{14}; (iii) the bias term β_i of the sensory neurons, of the inter-neurons, and of the neuron N_{14}. The decay constant τ_i of the sensory neurons and of the output neurons N_{12} and N_{13} is set to 0.1. Cell potentials are set to 0 any time the network is initialised or reset, and circuits are integrated using the forward Euler method with an integration step-size of 0.1.

A simple generational genetic algorithm is employed to set the parameters of the networks (see Goldberg, 1989). The population contains 80 genotypes. Generations following the first one are produced by a combination of selection with elitism, recombination and mutation. For each new generation, the three highest scoring individuals ("the elite") from the previous generation are retained unchanged. The remainder of the new population is generated by fitness-proportional selection (also known as roulette wheel selection) from the individuals of the old population. Each genotype is a vector comprising 84 real values (i.e., 70 connection weights, 6 decay constants, 7 bias terms, and a gain factor). Initially, a random population of vectors is generated by initialising each component of each genotype to values chosen uniformly random from the range [0,1]. New genotypes, except "the elite", are produced by applying recombination with a probability of 0.3 and mutation. Mutation entails that a random Gaussian offset is applied to each real-valued vector component encoded in the genotype, with a probability of 0.15. The mean of the Gaussian is 0, and its standard deviation is 0.1. During evolution, all vector component values are constrained to remain within the range [0,1]. Genotype parameters are linearly mapped to produce network parameters with the following ranges: biases $\beta_i \in [-4, -2]$ with $i \in \{1, 2, \ldots, 6\}$, biases $\beta_i \in [-5, 5]$ with $i \in \{7, 8, \ldots, 14\}$; weights $\omega_{ij} \in [-6, 6]$ with $i \in \{1, 2, \ldots, 6\}$ and $j \in \{7, 8, \ldots, 11\}$, weights $\omega_{ij} \in [-10, 10]$ with $i \in \{7, 8, \ldots, 11\}$ and $j \in \{7, 8, \ldots, 14\}$; gain factor $g \in [1, 13]$. Decay constants are firstly linearly mapped into the range $[-1.0, 1.3]$ and then exponentially mapped into $\tau_i \in [10^{-1.0}, 10^{1.3}]$. The lower bound of τ_i corresponds to the integration step-size used to update the controller; the upper bound, arbitrarily chosen, corresponds to about 1/20 of the maximum length of a trial (i.e., 400 s).

6.3.4 The fitness function

During evolution, each genotype is translated into a robot controller, and cloned onto each agent. Then, the group is evaluated twelve times, six in *Env. L*, and six in *Env. R*. The sequence order of environments within the twelve trials has no bearing on the overall performance of the group since each robot controller is reset at the beginning of each trial. Each trial (e) differs from the others in the initialisation of the random number generator, which influences the robots' starting position and orientation, and the noise added to motors and sensors. Within a trial, the robot life-span is 400 simulated seconds (4,000 simulation cycles). In each trial, the group is rewarded by an evaluation function f_e which seeks to assess the ability of the team to approach the light bulb, while avoiding collisions and staying within the range of the robots' infrared sensors:[1]

$$f_e = KP\Big(\sum_{t=i}^{T}[(d_t - D_{t-1})(\tanh(\frac{S_t}{\rho}))]\Big)$$

As in Quinn et al. (2003), the simulation timesteps are indexed by t and T is the index of the final timestep of the trial; d_t is the Euclidean distance between the group location at timestep t and its location at timestep $t = 0$, and D_{t-1} is the largest value that d_t has attained prior to timestep t. Therefore, the term $(d_t - D_{t-1})$ measures any gain that the team has made on its previous best distance from its initial location which is taken to be the centroid of the group.

The factor $\tanh(\frac{S_t}{\rho})$ reduces any fitness increment given by $(d_t - D_{t-1})$ when one or more robots are outside of the infrared sensor range: S_t is a measure of the team's dispersal beyond the infrared sensor range ρ ($\rho = 24.6$ cm) at timestep t. Recall that robot R_{AL} has no infrared sensors. Therefore, it does not have a direct feedback at each timestep of its distance from its group-mates. Nevertheless, the sound can be indirectly used by this robot to adjust its position within the group. If each robot is within ρ range of at least another, then $S_t = 0$. Otherwise, the two shortest lines that connect all three robots are found and S_t is the distance by which the longest of these exceeds ρ. The function $\tanh(x)$ assures that, as the robots begin to disperse, the team's score increment falls sharply.

$$P = 1 - (\sum_{i=1}^{3} c_i/c_{max}) \text{ if} \sum_{i=1}^{3} c_i \leq c_{max},$$

reduces the score in proportion to the number of collisions which have occurred during the trial, where c_i is the number of collisions of the robot i and $c_{max} = 4$ is the maximum number of collisions allowed. $P = 0$ if $\sum_{i=1}^{3} c_i > c_{max}$.

The team's accumulated score is multiplied by $K = 3.0$ if the group moved towards the light bulb, otherwise $K = 1.0$. Note that a trial is terminated early if (a) the team reached the light bulb (b) the team distance from the light bulb exceeds an arbitrary limit set to 150 cm, or (c) the team exceeds the maximum number of allowed collisions c_{max}.

---

[1] Note that, this fitness function is very similar to the one used in Quinn et al. (2003) from which it mainly differs for the parameter K. This parameter has been introduced to give a selective advantage to those groups which move towards the light bulb. In order to facilitate comparisons between our work and that detailed in Quinn et al. (2003), we provide a description of the fitness function which uses a similar mathematical notation employed in Quinn et al. (2003).

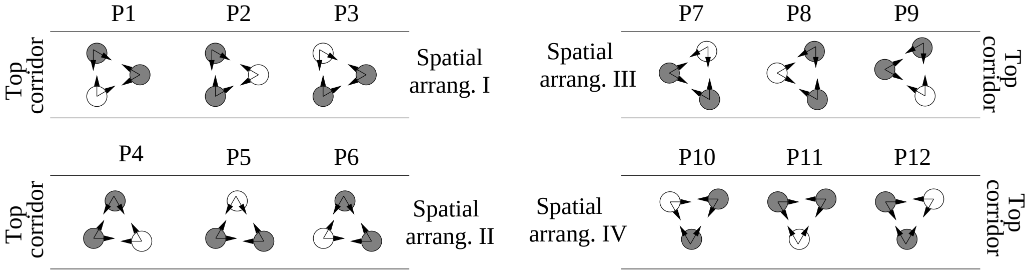


Figure 6.5: The robots' initial positions (from P1 to P12) during the post-evaluation phase. White circles refer to robot R_{AL}, gray circles refer to robot R_{IR}. For each robot, the black arrows indicate the region within which the robot's heading is randomly chosen. See text in section 6.4 for details.

6.4 Results

Ten evolutionary simulations, each using a different random initialisation, were run for between 2,500 and 3,600 generations of the evolutionary algorithm. The termination criterion for each run was set to a time equal to 86,400 seconds of CPU time. Given the way in which the fitness is computed and the dimensions of the world, scores higher than 200 refer to groups that manage to repeatedly get very close to the light in both types of environment. Several runs produced successful groups, however the fitness of the best groups of the most successful evolutionary runs oscillates quite a lot throughout the evolution. These oscillations might be related to the phenomenon of the overestimation of the fitness of the best groups. Maybe, during evolution, the best groups took advantage of favourable conditions, which are determined by the existence of between-generation variation in the starting positions and relative orientation of the robots and other simulation parameters. Thus, in the next section, we show the results of a first series of post-evaluation tests aimed to estimate the effectiveness of the best evolved navigation strategies of each run, under circumstances in which the effect of favourable conditions linked to the initialisation of the robots are ruled out.

6.4.1 First post-evaluation tests

In order to have a better estimate of the behavioural capabilities of the evolved controllers, we post-evaluated, for each of the 10 evolutionary simulations, the genotype with the highest fitness. The groups of robots controlled by neural networks built from these genotypes are referred to as g_1 to g_{10}. During post-evaluation, each group is subject to a set of 1,200 trials in both environments. The number of post-evaluation trials per type of environment (i.e., 1,200) is given by systematically varying the initial positions of the three robots according to the following criteria: (i) we defined four different types of spatial arrangements in which the robots are placed at the vertices of an imaginary equilateral triangle inscribed in a circle of radius 12 cm and centred in the middle of the top corridor (see figure 6.5); (ii) for each spatial arrangement, we identified three possible relative positions of the robot R_{AL} with respect to the walls of the corridor (see white circle in figure 6.5); (iii) for each of these (four times three) initial positions, the post-evaluation is repeated one hundred times. The initial orientation of each robot is determined by applying an angular displacement randomly chosen in the interval $[-30°, 30°]$ with respect to a vector originating from the centre of the robot and pointing towards the centroid of the group. The four times three different arrangements take into account a set of relative positions among the robots and between the robots and the walls so that the success rate of the group is not biased by these elements.

We decided to estimate the effectiveness of the robots' behavioural capabilities during post-evaluation by employing a binary criterion (successful/unsuccessful) instead of the fitness function as during evolution. In particular, a group is considered successful if its centroid is less than 10 cm away from the light bulb. However, preliminary tests showed that the satisfaction of this criterion in 400 sec was too demanding for the robots. Many of the initial positions resulting from the systematic variation as explained above require the robots to spend a lot of the time at their disposal in re-arranging themselves to be able to safely progress towards the light, leaving little time for navigation. It appeared that some of the evolutionary conditions (e.g., the random initialisation of the robots' initial position and the few evaluation trials) did not favour groups capable of quickly arranging themselves for phototaxis regardless their initial positions. Consequently, even groups capable of moving towards the light without colliding resulted very often unsuccessful due to lack of time to fulfil the criterion mentioned above. Since our interest is on collision free navigation strategies and not on other characteristics of the phototactic movement such as the speed (i.e., how quickly the robots get to the light bulb), we decided to make the post-evaluation trials 2.5 times longer than the trials during evolution (i.e., 1,000 s, 10,000 simulation cycles). This should (i) give the robots enough time to compensate for possible disruptive effects induced by initial positions never or very rarely experienced during evolution, and (ii) provide us with a fair estimation of the navigation capabilities of each of the groups selected for post-evaluation. At the beginning of each post-evaluation trial, the controllers are reset (see section 6.3.3 for details).

The results of the post-evaluation phase are shown in figure 6.6. We notice that the best groups are g_9 and g_{10}, that achieve a performance over 90% in both environments. Groups g_4 and g_7 display a performance over 80% in both environments. The performance of all the other groups is clearly unsatisfactory. Groups g_2, g_3, g_5 and g_8 proved to be capable of accomplishing the task only when located in an *Env. R*, while group g_1 is particularly effective in *Env. L*. This phenomenon can be explained by considering that the two environments require two different types of turn—a left turn in *Env. L*, and a right turn in *Env. R*. By looking at the behaviour of the groups through a simple graphical interface, we observed that the successful groups employ two different navigation strategies to make the two types of turn (see section 6.4.2). We also observed that those groups that systematically fail in any of the two environments, lack the capability to make both turns. Note that when looking at the performances of the best evolved groups, as shown in figure 6.6, one has to take into account the arbitrary criteria we chose to determine whether or not a group of robots is successful in any given trial: no robot has to collide with the walls or with the other robots. This is a very strict condition, which, given the nature of the task, demands each agent to be very accurate in coordinating its movement. Some

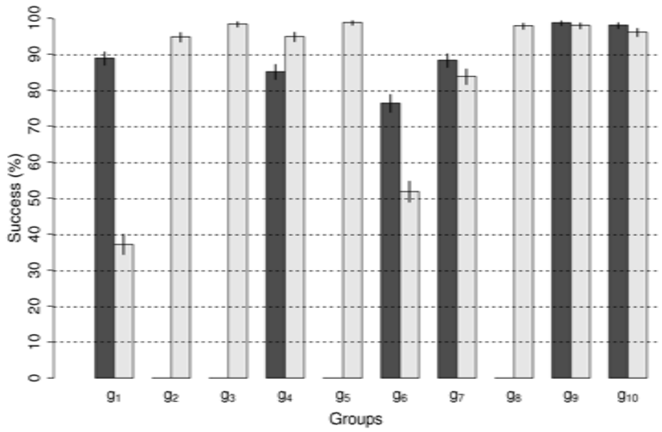


Figure 6.6: Results of post-evaluation showing the success rate (%) with confidence interval (computed with the binomial test) over 1,200 trials per type of environment (black bars refer to *Env. L*, and white bars to *Env. R*) of the groups of robots (g_1 to g_{10}) whose controllers are built from the genotype with the highest fitness of each evolutionary simulation.

initial tests indicate that if we allow the group to make a certain number of collisions (e.g., four) before defining a trial as a failure, then several groups would result almost always successful in both types of environment (data not shown). Whether or not the robots should be allowed to collide or the extent to which a single collision invalidates the performance of the group are issues that go beyond the scope of this research work and shall not be discussed any further. Instead, we focus on other performance measures which tell us more about the characteristics of the best evolved groups. For instance, by looking at the data shown in table 6.1, we notice that, except for group g_2, the majority of the failures in *Env. L*, are due to collisions. In *Env. R*, the performances of all the groups, are sensibly better than those in *Env. L* (see columns 4 and 5, table 6.1). If we look at the average distances to the light (see columns 6 and 7, table 6.1) and the relative standard deviations (see columns 8 and 9, table 6.1), we can see that in *Env. L* failures happen rather far away from the light. For example, for groups g_3, g_5 and g_8—100% unsuccessful in *Env. L*—the final distance to the light is almost equal to the initial distance. This denotes a lack of coordination of movement during the initial phase, when the robots have to assume a configuration which favours the group phototaxis. In *Env. R*, the smaller final distances to the light seem to denote a problem, possibly common to several groups, in making the right turn.

In the rest of this section, we concentrate on the analysis of the group g_9, which proved to be the most effective in the first post-evaluation test. The tests we are going to illustrate have been carried out for all the best evolved groups. It turned out that, successful navigation strategies of any best evolved group are very similar from a behavioural point of view, and in terms of the communication mechanisms exploited to obtain the coordination of actions. Therefore, the reader should consider the operational description of the behaviour of group g_9 representative of all the successful navigation strategies of any best evolved group. These groups seems to differ in terms of the robustness of the mechanisms that underpin their behaviour rather than on the nature of these mechanisms.

6.4.2 A description of the behavioural strategies

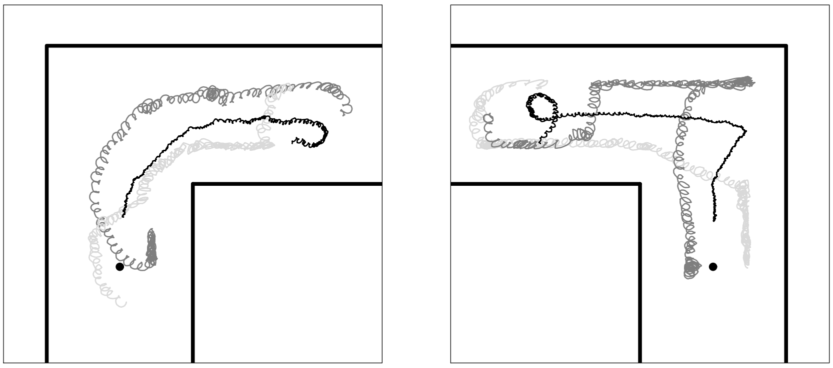
In this section we provide a qualitative description of the individual motion of robots of group g_9 as observed through a simple graphical interface. First of all, we noticed that the systematic variation of the initial positions of the robots during post-evaluation brings about contingencies in which the coordination of movements of the group toward the target requires an initial effort of the robots in re-arranging their relative positions.[2] During this initial phase of a trial a dynamic process guided by the nature of the flow of sensations induces the specialisation of the controllers with respect to the physical characteristics of the robots, and to the relative role that they play in the group. This phase is followed by the navigation phase in which the group maintains a rather regular spatial configuration; that is, the two robots R_{IR} place themselves in between the target and the robot R_{AL}. However, note that while *Env. L* requires the group to make a left turn, *Env. R* requires the group to make a right turn. This asymmetry in the environmental structures corresponds to differences in behavioural strategies employed by the group to reach the target as shown

Table 6.1: Further results of the post-evaluation test, showing for the best evolved groups: (i) the percentage of unsuccessful trials due to exceeded time limit without the group having reached the target (columns 2, and 3); (ii) the percentage of unsuccessful trials which terminated due to collisions (columns 4, and 5); (iii) the average and standard deviation of the final distance of the centroid of the group to the light during the unsuccessful trials (respectively columns 6, 8 for *Env. L*, and columns 7, 9 for *Env. R*). Note that in all trials the initial distance between the centroid of the group and the light is equal to 85.14 cm.

	(%) of failure due to time limit		**(%) of failure due to collisions**		**Distance to the light**			
					avg		**std**	
group	***Env. L***	***Env. R***	***Env. L***	***Env. R***	***Env. L***	***Env. R***	***Env. L***	***Env. R***
g_1	0.00	52.75	10.92	9.92	82.19	52.17	6.63	32.74
g_2	85.33	1.83	14.67	3.17	66.30	46.17	4.36	18.22
g_3	0.00	1.00	100.00	0.42	81.02	36.38	4.049	12.02
g_4	0.67	0.50	14.00	4.50	57.83	69.30	13.50	15.32
g_5	0.00	0.00	100.00	1.00	79.05	41.13	2.94	12.93
g_6	0.00	31.00	23.42	17.08	77.50	64.27	11.89	29.92
g_7	0.58	10.00	10.92	6.00	50.98	40.80	30.18	14.18
g_8	0.00	0.00	100.00	1.92	80.94	53.03	2.34	11.59
g_9	0.00	0.83	1.08	1.00	77.71	50.22	11.44	21.08
g_{10}	0.00	2.17	1.75	1.50	82.28	90.37	13.19	31.81

---

[2] The movies of the performances of the group in both environments are available at `http://iridia.ulb.ac.be/supp/IridiaSupp2006-006/`.



(a) (b)

Figure 6.7: Trajectories of the agents during a successful trial in (a) *Env. L*; (b) *Env. R*. The black lines refer to the trajectories of robot R_{AL} while the other lines refer to the trajectories of robots R_{IR}. The thick horizontal and vertical segments represent the walls. In each figure, we depict only the side of the corridor where the light (the small black dot) is located.

in figure 6.7. While in *Env. L* the robots simply turn towards the light keeping their relative positions in the group, in *Env. R* we first observe an alignment of the agents along the far right wall (see figure 6.7b). Subsequently, the agent close to the corner (see the dark gray line) overcomes the other two and the group starts moving towards the target once the classical configuration of the two robots R_{IR} in between the target and the robot R_{AL} is re-established.

Another important qualitative element is that each of the members of the group is characterised by a movement with a strong angular component (anti-clockwise). In other words, the robots proceed toward the light by rotating on the spot. Within a trial, pure linear movement replaces the rotational behaviour only sporadically and for a very short interval. This can happen to avoid an imminent danger of collision or if required by the navigational strategy of the group. The evolution of the rotational movement is not particularly surprising if we think about its effect on the perception of sound. In particular, the rotational movement can introduce rhythm in perception. The oscillations of perceived sound, produced by the rotational movement and/or by the oscillations manifested in signalling behaviour, can provide the robots the cues to adjust their relative positions. Further and deeper investigations on the nature of sound signals and its relationship with the robots' motion will be carried out in the next sections.

The effect of the starting position and the rotational movement are phenomena that have a strong effect on the time it takes the group to reach the target. Indeed, as resulted from the post-evaluation test, most of the successful trials of group g_9 last longer than the 400 sec given to the groups to complete the task during the evolutionary phase (data not shown).

6.4.3 A description of the signalling behaviour

The group is required to coordinate its actions to maintain its cohesion without incurring into collisions and to get closer to the target. What is the role of signalling for the achievement of these goals? Is signalling used by robot R_{AL} to communicate to robots R_{IR} information concerning the relative position of the target? Similarly, is signalling used by robots R_{IR} to inform robot R_{AL} on the position of obstacles against which it may collide? In order to provide an answer to this type of questions, we carried out a series of tests that look at the properties of the sound signals perceived by each robot during a successful trial in each environment. Our goal is to identify oscillatory phenomena or other distinctive features in sound production/perception whose properties can be exploited by the robots to coordinate their actions. Despite the fact that our analysis is limited to two successful trials, one for each type of environment, we hope that the results help us formulate general hypotheses concerning the role of signalling in the coordination of the group.

Before proceeding further, we should remind the reader that the intensity of sound perceived at each microphone results from the summation of two components (the self and the non-self) and the noise. The self component (i.e., the agent's own signal) is only determined by the intensity of the sound emitted by the robot itself. The non-self component is determined by the intensity at which the sound is emitted from the loudspeaker of a sender as well as by the relative distance and orientation of the loudspeaker with respect to the receiver's microphones (see section 6.3.2). Although the agents have no means to distinguish between the self and the non-self components of the perceived sound, they can act in a way to determine patterns in the flow of sensations which are informative on their spatial relationship.

For robots of group g_9, we proceeded by separately recording the self and the non-self components of the sound perceived at each microphone, and the heading at each timestep of each robot during a successful trial in each environment. As mentioned in section 6.4.2, each robot of group g_9 combines phototaxis with a rotational movement. The latter, due to the simulated physics, can introduce rhythms in the perception of sound through its effects on the non-self component. With a Fast Fourier Transform analysis (FFT), we transform the sequences of heading and the self and non-self components of the sound signal perceived by each robot at each microphone from the time domain into the frequency domain. By looking at the power spectral density (PSD) we observe that: (a) the self component of each robot does not display any harmonic (fn_i) at any frequency different from 0 Hz; (b) for all the robots, there are three principal harmonic components in the spectrum of the sequence of heading (see table 6.2 columns 3, 4, 5); (c) the non-self component of each robot has only one principal harmonic (see table 6.2 columns 6, and 7); (d) fn_1 of robot R_{AL} differs from fn_1 of both robots R_{IR}.

From point (a) we conclude that for each robot there are no oscillatory phenomena in sound production. Oscillations are instead observed in the perceived sound. From points (a), (b) and (c) we conclude that oscillations of the perceived sound are produced by

Table 6.2: Frequencies (Hz) of the principal harmonic components fn_i in the sequence of heading (columns 3, 4 and 5), and in the non-self component of the perceived sound signal at sensor S_1 (columns 6) and at sensor S_2 (columns 7) for each robot.

		heading			**non-self**	
					S_1	S_2
Env.	**robot**	fn_1 **(Hz)**	fn_2 **(Hz)**	fn_3 **(Hz)**	fn_1 **(Hz)**	fn_1 **(Hz)**
Env. L	R^1_{IR}	0.187	0.375	0.566	0.185	0.185
	R^2_{IR}	0.187	0.377	0.556	0.181	0.181
	R_{AL}	0.205	0.411	0.617	0.205	0.205
Env. R	R^1_{IR}	0.186	0.378	0.563	0.184	0.184
	R^2_{IR}	0.184	0.358	0.559	0.181	0.181
	R_{AL}	0.212	0.426	0.633	0.201	0.201

the rotational movement of each robot through the effect that the movement has on the characteristic of the non-self components. A further evidence of the causal relationship between the rotational movement and the oscillation of the non-self components is given by the fact that the principal harmonic of the non-self components has a very similar frequency to the first harmonic of each robot's sequence of heading (see table 6.2, columns 3 and 6). Moreover, the similarities of the first harmonic of the non-self component between robots R_{IR} and the differences between robots R_{IR} and robot R_{AL} confirm that there is a dynamic speciation of the characteristics of the homogeneous controllers with respect to the physical properties of the robots. In particular, robot R_{AL} rotates slightly faster than the other two robots.

So far, we have identified periodic phenomena and their relative frequencies in sound signals and in the rotational movement of the robots. The next step of our analysis focuses on the characteristics of the non-self components. We use the frequencies of the principal harmonic fn_1 obtained from the PSD analysis to filter the sound signals. In particular, we applied a narrow bandpass filter at frequencies 0 Hz and fn_1. In this way, we transform the non-self components into sinusoidal signals $ns_i = \alpha + p\sin(2\pi fn_1 t)$, where α is the DC offset[3] of the signal, p is the peak amplitude, and fn_1 is the frequency of oscillation. From our analysis, it results that the average amplitude and standard deviation of the self components and the DC offset and peak of ns_i with $i \in \{1, 2, 3\}$ do not substantially differ (i) among the robots; (ii) between the two sensors (S_1 and S_2); and (iii) between the two environments (i.e., *Env. L* and *Env. R*).

In particular, the mean value of the self components contributes to more than 90% of the perceived sound (see table 6.3 columns 2, and 8).[4] Given the high intensity of the self component, the non-self component can only induce changes in the perception of sound that are less than 10% of the sensors' receptive-field. By looking at the DC offset of the sinusoidal signals ns_i (see table 6.3 columns 4, 6, 10 and 12) we can infer that the non-self

---

[3] DC offset is the mean amplitude of a waveform; if the mean amplitude is zero, there is no DC offset.

[4] The self components are described by referring to their average and standard deviation since they do not present any periodic oscillations.

Table 6.3: This table shows, for each robot of group g_9 and for each microphone (S_1, and S_2): (i) the average and standard deviation of the intensity of the self component (columns 2, 3, 8, and 9); (ii) the DC offset (columns 4, 6, 10, and 12) and peak (columns 5, 7, 11, and 13) of the filtered non-self component ns_i. Data refers to sound intensity values recorded during a successful trial in each type of environment. See text in section 6.4.3 for details on the filtering.

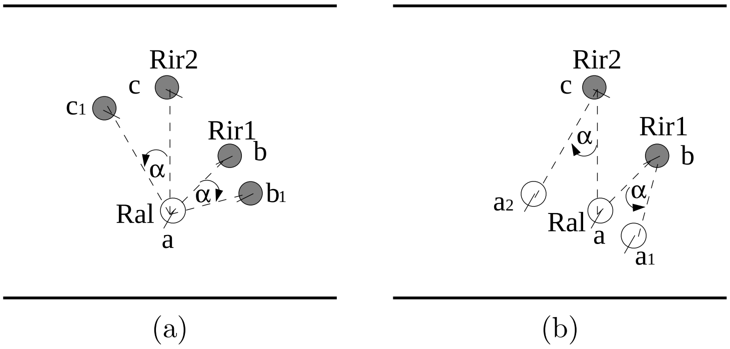
	Env. L						***Env. R***					
	self		**sinusoidal signal $\mathbf{ns_i}$**				**self**		**sinusoidal signal $\mathbf{ns_i}$**			
			S_1		S_2				S_1		S_2	
	avg	**std**	**DC**	**peak**	**DC**	**peak**	**avg**	**std**	**DC**	**peak**	**DC**	**peak**
R^1_{IR}	0.934	0.030	0.115	0.076	0.106	0.056	0.937	0.027	0.103	0.070	0.093	0.051
R^2_{IR}	0.932	0.035	0.133	0.018	0.120	0.014	0.933	0.035	0.107	0.054	0.098	0.041
R_{AL}	0.926	0.018	0.115	0.064	0.115	0.060	0.921	0.020	0.135	0.014	0.136	0.012

components are already very weak, possibly due to the relatively far robot-robot distances. Despite this, if we sum, for each robot (i), for each sensor (s) and for each environment, the average intensity of the self component, the DC offset and the peak of the ns_i, we obtain values that give us an indication of what could be the reading of the sound sensors when the non-self components are at their highest intensity. Since these values are higher than 1, it follows that the reading of the sound sensors saturate (i.e., $\hat{S}_{is} = 1$). From this we infer that, if not attenuated by the shadowing effect, the non-self plus the self component can be sufficient to saturate the sensors' receptive field of the receiver. If we combine this data with the fact that the non-self components oscillate due to the rotational behaviour of the robots, we can formulate the following hypothesis: during navigation, the readings of the sound sensors of each robot may go through oscillations constrained between an upper and a lower bound. The upper bound is reached when the sum of the self and the non-self component corresponds to a value equal or bigger than the saturation value of the sound sensors (1.0). The lower bound is close to the intensity of the self component that is reached when the non-self components are strongly attenuated by the shadowing effect.

These oscillations are very small since they concern less than 10% of the auditory receptive field, and certainly not very regular since the random noise applied to the sensors reading can disrupt the regularity of the oscillations determined by the contingencies (i.e., rotational movements and robots' relative distances). However, in spite of being small and noisy, these oscillations seem to be the only phenomenon related to the perception of sound that play a significant role in the coordination of the group. In fact, given a controller sufficiently sensitive to capture them, they may represent a valuable perceptive cue for the receiver to spatially discriminate sound sources and consequently relative position and orientation of the emitter/s. For example, low intensity of sound corresponds to conditions in which the body of the receiver is placed in between its sound receptor and the sound source; high intensity of sound corresponds to conditions in which the sound receptor is in between the body of the receiver and the sound source. Robots capable of detecting these spatial relationships can use them to make movements towards or away from a sound source. Moreover, the oscillations of perceived sound, produced by the rotational movement might emphasise the intensity differences between the two sound receptors. These differences also known as Interaural Intensity Differences (hereafter referred to as IIDs, see Kandel et al., 2000) could provide the robots the cues to adjust their positions relative to each other. These cues might be exploited by the robots to remain close to each other while avoiding collisions and moving towards the target. Given the lack of complexity in robots' sound production, we exclude that signalling behaviour concerns more articulated forms of communication. In the following section, we show the results of further post-evaluation tests that are meant to assess whether oscillations in sound perception and IIDs are cues used by the robots for spatial discrimination of sound sources and coordination of actions.

6.4.4 Signalling behaviour and the group's coordination of actions

In this section, our goal is to find out whether both types of robots exploit sound signalling to mutually coordinate their actions. Alternatively, the use of sound may be limited to robot R_{AL}. Robots R_{IR} may ignore the sound and base their movements on the readings of the infrared sensors. This would be sufficient to keep both robots R_{IR} close to robot R_{AL}. The latter, by moving towards the target, will inevitably bring the group to the light. Another possibility is that none of the robots use sound. In this case, the group might employ unchanging phototactic movement which may work as well given that the dimensions of the corridors and the positions of the lights in the two environments do not vary. For example, robot R_{AL} may move for about 65 cm east/west according to the



(a) (b)

Figure 6.8: (a) *Test A*: robots R_{IR} (the gray circles) are displaced of an angle α with respect to robot R_{AL} (empty circle). This picture represents a hypothetical state in which the readings of the sound sensors of robot R_{AL} are computed considering R^1_{IR} located in position b_1 instead of b, and R^2_{IR} located in position c_1 instead of c. (b) *Test B*: robot R_{AL} is displaced of an angle α with respect to robots R_{IR}. This picture represents a hypothetical state in which the readings of the sound sensors of robot R^1_{IR} are computed considering R_{AL} located in position a_1 instead of a. The sound sensors of robot R^2_{IR} are computed considering R_{AL} located in position a_2 instead of a.

characteristics of the environment and then south; robots R_{IR} have simply to follow R_{AL} avoiding collisions.

In the following, we run two post-evaluation tests, *Test A*, and *Test B*. In both tests, we interfere with the propagation of sound in the environment by disrupting the orientation of the robot emitter with respect to the heading of the receiver (see figure 6.8). In particular, in each test, the robots undergo sets of 1,200 trials in each type of environment. For all the simulation cycles following the first 10 seconds[5] of each trial of a set, the sound sensors reading of a type of robot (i.e., R_{AL} or R_{IR}) are computed with respect to a hypothetical state of the system in which each robot of the other type is supposed to be re-oriented by a fixed angular displacement, ranging from a minimum of 18° to a maximum of 180°, with a randomly chosen direction (clockwise or anti-clockwise) with respect to the heading of the receiver. The magnitude of the angular displacement does not vary in each set of 1,200 trials in a given environment. Note that, the updating of the infrared sensors of robots R_{IR} and of the ambient light sensors of R_{AL} do not undergo any disruption during these tests. The hypothetical states are taken into account only as far as it concerns the updating of the sound sensors' reading of one type of robot at the time. In particular, in *Test A*, the sound perceived by robot R_{AL} is computed with reference to a hypothetical state in which the orientation of both robots R_{IR} with respect to R_{AL}'s heading is changed in order to meet the angular displacement requirements (see figure 6.8a). No disruptions are applied to update the sound perceived by robots R_{IR}. In *Test B*, the sound perceived by the robots R_{IR} is computed with reference to a hypothetical state in which the orientation of robot R_{AL} with respect to the R_{IR}'s heading is changed in order to meet the angular displacement requirements (see figure 6.8b). In this type of tests no disruptions are applied to update the sound perceived by robot R_{AL}.

By varying the sender-receiver orientation, we indirectly increase/decrease the magnitude of the non-self component. In particular, those hypothetical states which tend to shorten the length of the shadowed path of the non-self components with respect to what indicated by the status of the system without disruptions, produce an increase of the intensity of the non-self component (see section 6.3.2). Consequently, these circumstances induce an increase of the proportion of time in a trial the sound sensors are saturated. On the contrary, those hypothetical states which tend to increase the length of the shadowed path of the non-self components with respect to what indicated by the status of the system without disruptions, produce a decrease of the intensity of the non-self component (see section 6.3.2). Consequently, these circumstances induce a decrease of the proportion of time in a trial the sound sensors are saturated.

From what said above, we can infer that *Test A* and *Test B* disrupt any kind of regularities in the perception of sound which are linked to sender-receiver relative orientation. In particular, in section 6.4.3, we have seen that oscillations of the perceived sound and IIDs are the only two phenomena of signalling behaviour which might be used by the robots to coordinate their actions. In *Test A* and *Test B*, spatial cues provided by these two phenomena do not refer anymore to the current status of the system but to hypothet-

---

[5] Applying any disruptions after 10 sec (i.e., 100 simulation cycles) gives the controllers sufficient time to reach a functional state different from the initial one, arbitrarily chosen by the experimenter, in which the cell potential of the neurons is set to 0 (see section 6.3.3).

ical states artificially induced. Consequently, a drop in the group performance at *Test A* is a sign that these cues are exploited by the robot R_{AL} to successfully carry out their task. Similarly, a drop in the group performance at *Test B* is a sign that these cues are exploited by the robots R_{IR} to successfully carry out their task. If both tests show a drop in the group performance, we would say that sound signalling is a common means of communication exploited by both types of robots to mutually coordinate their actions.

The results of *Test A* are shown in figure 6.9a and 6.9b. The results of *Test B* are shown in figure 6.9c and 6.9d. From these graphs we notice that the performance of the group is significantly disrupted by alterations which concern the orientation of one type of robot with respect to the heading of the other type of robot. In particular, the bigger the magnitude of the angular displacement, the higher the percentage of failure of the system. The majority of failures are due to robot-wall collisions. Observing the behaviour of the group in these conditions, we noticed that, under the effects induced by the disruptions, the robots are not capable of remaining close to each other—i.e., within the infrared sensors' range. When the distances becomes too high, the robots start wandering around the arena, and the trial terminates due to a collision of the robot R_{AL} with the arena walls. Only in few circumstances the robots do not lose contact to each other but they are not capable of reaching the target within the time limits (see figure 6.9 black area of the bars).

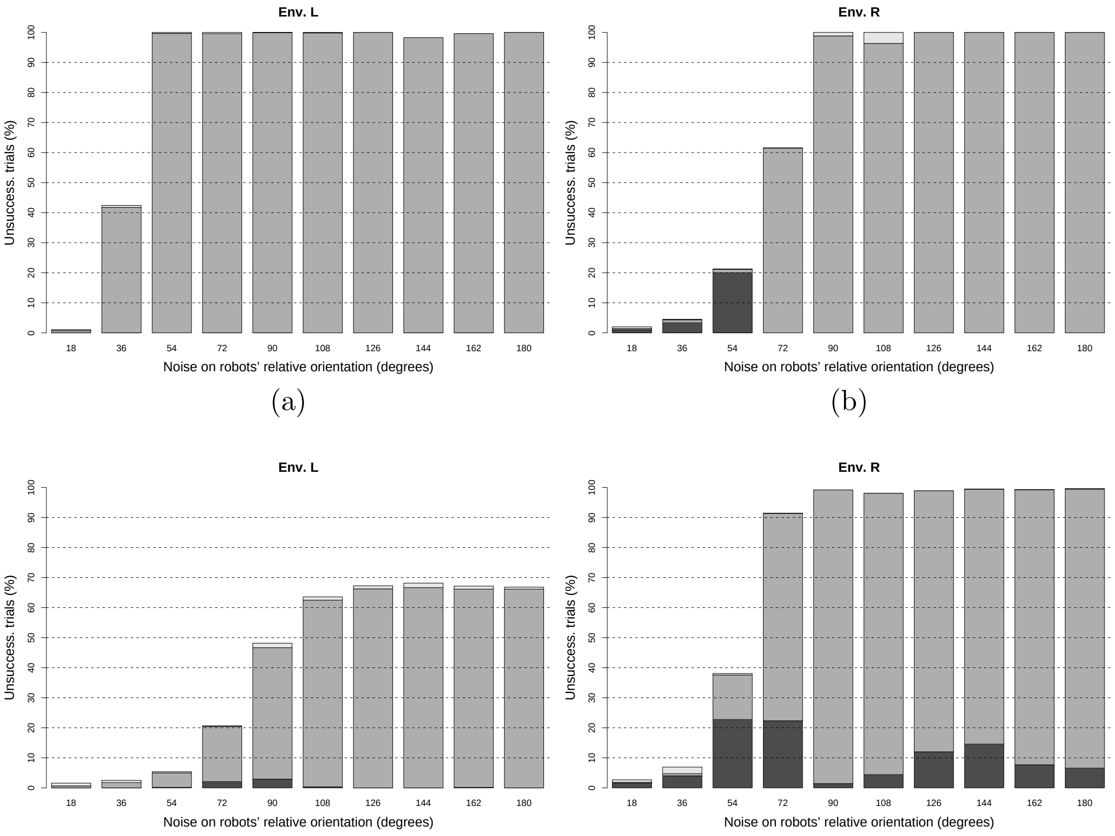
These results prove that the group performance is severely disrupted when the hypothetical status of the system, used to update the sound sensor readings of either type of robot, is significantly different from the current circumstances. If the oscillations of the sound sensors' reading and IIDs of either type of robots do not reflect the environmental contingencies, the group performance in both environments is disrupted. We conclude that, for both types of robots, spatial cues provided by the oscillations of perceived sound and possibly by IIDs have a bearing on the development of effective navigational strategies. Sound signalling seems to be a common means of communication exploited by both types of robots to mutually coordinate their actions.

6.4.5 The significance of the Interaural Intensity Differences (IIDs)

In this section, we illustrate the results of further post-evaluation tests whose goal is to determine whether IIDs are cues used by the robots to coordinate their actions. Alternatively, oscillations of perceived sound may be sufficiently informative on the contingencies to allow the robots to successfully accomplish their goal.

In these tests we progressively reduce the IIDs up to the point at which the two sound receptors (S_1 and S_2) of a type of robot are impinged by the same stimulus. Consequently, these disruptions hinder the possibility of the robots to use IIDs as cues for localisation of sound sources and for their coordination of actions. At the same time, we preserve the phenomenon of oscillations of perceived sound as cues for spatial discrimination and localisation of sound sources.

In each test, the robots undergo sets of 1,200 trials in each type of environment. For all the simulation cycles following the first 10 seconds of each trial of a set, the reading of one sound sensor (i.e., S_1 or S_2) of a type of robot (i.e., R_{AL} or R_{IR}) are modified

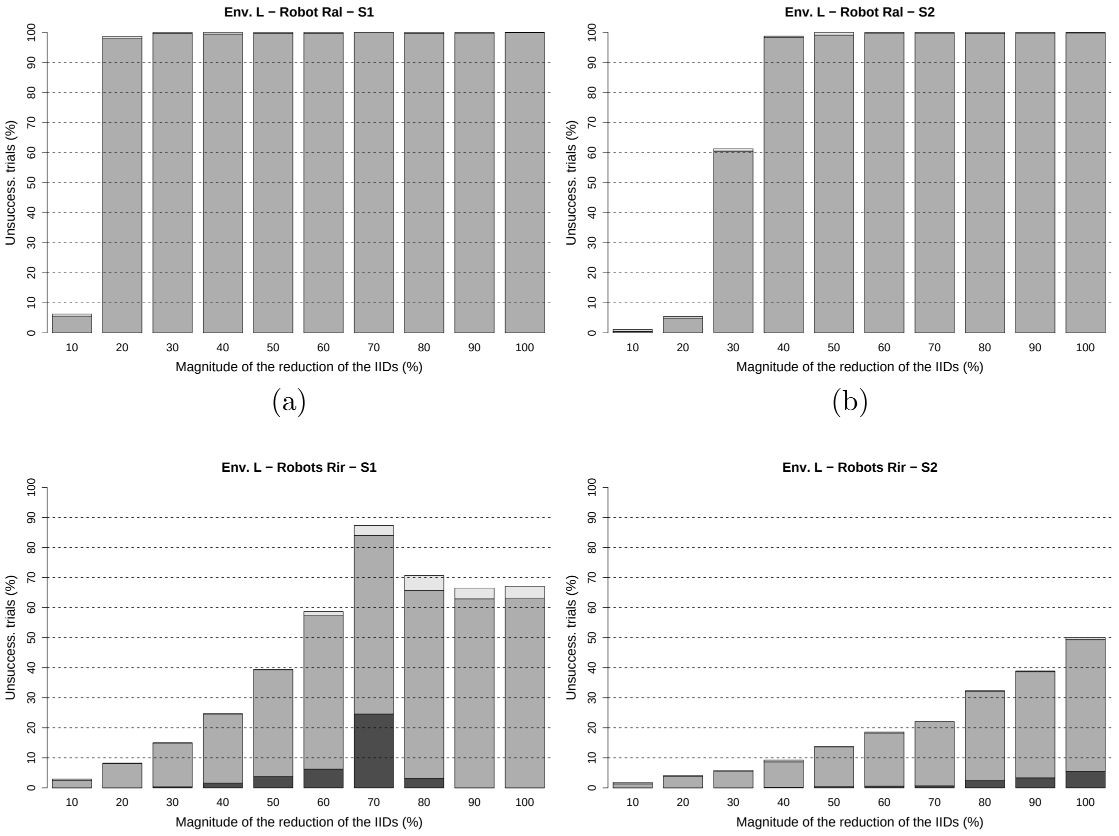


(a) (b)

(c) (d)

Figure 6.9: Percentage of failure during 1,200 trials in each type of environment in post-evaluation tests with disruptions applied to the relative orientation of the robots during the computation of the perceived sound. (a) and (b) refer to *Test A*. The robots R_{IR}, during all the simulation cycles following the first 10 seconds of any trial, are considered to be re-oriented with respect to the heading of robot R_{AL} by applying the angular displacement indicated on the horizontal axis and randomly choosing the direction of displacement (i.e., clockwise or anti-clockwise). (c) and (d) refer to *Test B*. The robot R_{AL} is re-oriented with respect to the heading of each robot R_{IR} as explained above. (a) and (c) refer to tests in *Env. L*; (b) and (d) refer to tests in *Env. R*. The black area of the bars refers to the percentage of trials terminated without collisions and with the group not having reached the target. The light gray area of the bars refers to the percentage of trials terminated due to robot-robot collisions. The dark gray area of the bars refers to the percentage of trials terminated due to robot-wall collisions.

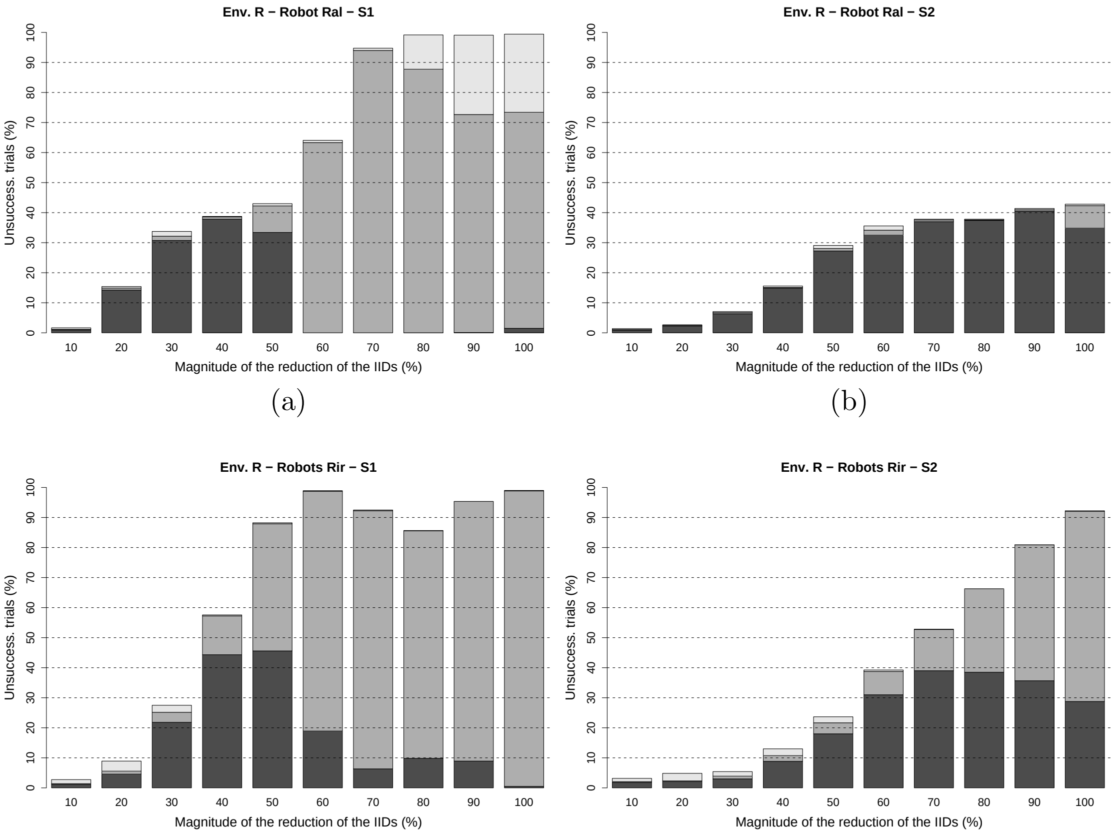
in order to reduce the IID of a given percentage, ranging from a minimum of 10% to a maximum of 100% (i.e., both sensors return the same reading). The magnitude of the decrease of the IIDs does not vary in each set of 1,200 trials in a given environment. Disruptions which reduce the IIDs of a given percentage are independently applied to (i)



(a) (b) (c) (d)

Figure 6.10: The graphs show the percentage of failure during 1,200 trials in *Env. L* with disruptions applied to: (a) robot R_{AL} sound sensor S_1; (b) robot R_{AL} sound sensor S_2; (c) robot R_{IR} sound sensor S_1; (d) robot R_{IR} sound sensor S_2. The disruptions concern the decrease of the the IIDs of the percentage indicated on the horizontal axis. The black area of the bars refers to the percentage of trials terminated without collisions and with the group not having reached the target. The light gray area of the bars refers to the percentage of trials terminated due to robot-robot collisions. The dark gray area of the bars refers to the percentage of trials terminated due to robot-wall collisions.

robot R_{AL} and robots R_{IR}, (ii) sound sensor S_1 and S_2, and (iii) *Env. L* and *Env. R*, for a total of 8 different types of tests—two types of robots times two types of sound sensor times two types of environment. If we observe a sensible drop of the percentage of success of the group for disruptions applied to any of the two types of robot in any of the two types of environment regardless of the sound sensor disrupted (S_1 or S_2), then we conclude that, for that type of robot, IIDs are cues used to coordinate its actions during the navigation towards the light. In any other circumstances, we conclude that the oscillations of perceived sound, without IIDs, are sufficiently informative on the contingencies to allow



(a) (b)

(c) (d)

Figure 6.11: The graphs show the percentage of failure during 1,200 trials in *Env. R* with disruptions applied to: (a) robot R_{AL} sound sensor S_1; (b) robot R_{AL} sound sensor S_2; (c) robot R_{IR} sound sensor S_1; (d) robot R_{IR} sound sensor S_2. See caption of figure 6.10 for details.

the robots to successfully accomplish their goal.

The results of the full series of tests show that for both types of robots and for both sound sensors, the progressive reduction of the IIDs is associated with a drop in performance of the group. When disruptions are applied to robot R_{AL} in *Env. L* (see figure 6.10a and b) and of robots R_{IR} in *Env. R* (see figure 6.11c and d), regardless the sound sensor disrupted, the rate of failure of the group in a type of environment is above 90% when the IIDs are made unavailable to a type of robot (see figure 6.10a and b and figure 6.11c and d, last bar of each graph). In these cases, we conclude that, IIDs are cues strictly necessary for a specific type of robot to be able to coordinate its actions.

In all the other cases (robot R_{AL} in *Env. R*, and robots R_{IR} in *Env. L*), although a progressive reduction of the IIDs corresponds to a drop in performance of the group, the disruptions applied to any of the two sound sensors do not equally affect the performance of the group. With the IIDs completely removed, the rate of failure of the group range in

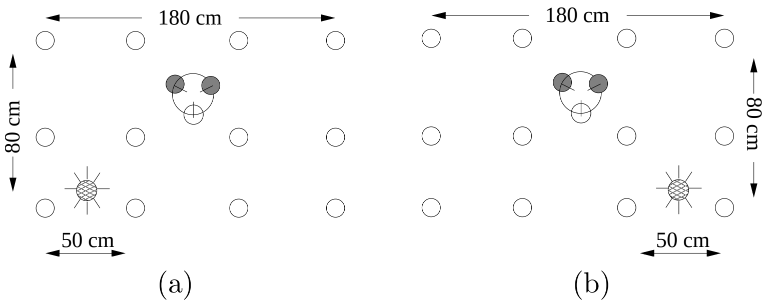
Table 6.4: Results of two post-evaluation tests for group g_9. In *Test C* the group is evaluated in a boundless arena; in *Test D* the group is located in two different types of environment as depicted in figure 6.12. The last row, in gray, shows the results of the post-evaluation test without disruptions in *Env. L* and *Env. R.* Columns 2, and 3 show the percentage of unsuccessful trials due to exceeded time limit without the group having reached the target. Columns 4, and 5 show the percentage of unsuccessful trials which terminated due to collisions. Columns 6, 7, 8, and 9 show the average and standard deviation of the final distance of the centroid of the group to the light during the unsuccessful trials (respectively columns 6, 8 for the environment that requires a left turn, and columns 7, 9 for the environment that requires a right turn). Note that in all trials the initial distance between the centroid of the group and the light is equal to 85.14 cm.

	(%) of failure due to time limit		**(%) of failure due to collisions**		**Distance to the light**			
					avg		**std**	
Test	**left turn**	**right turn**	**left turn**	**right turn**	**left turn**	**right turn**	**left turn**	**right turn**
C	76.50	81.75	0.00	0.00	20.73	20.34	2.54	2.27
D	15.33	10.00	16.25	1.67	29.49	28.76	7.09	12.01
	0.00	0.83	1.08	1.00	77.71	50.22	11.44	21.08

between 40% and 60% in case in which disruptions concern (i) robot R_{AL} sound sensor S_2 in *Env. R*; and (ii) robots R_{IR} in *Env. R* (see figure 6.11). These results suggest that the group is able to assume spatial configurations which facilitate the navigation in spite of the absence of IIDs for either type of robot.

6.4.6 Robustness to environmental changes

In this section, we show the results of final post-evaluation analyses that aim to test the robustness of the group navigation strategy in environments that differ from those experienced by the robots during the evolutionary phase. Contrary to the evolution, in *Test C*, we simply remove the walls. The absence of walls causes a sensible increase in the percent-



(a) (b)

Figure 6.12: Environment that requires (a) a left turn; (b) a right turn. These environments differ from *Env. L* and *Env. R* because the walls are replaced by round obstacles of the diameter of a robot.

age of failure in both environments (see columns 2, 3, table 6.4). The magnitude of the disruption suggests that the walls are environmental structures systematically exploited by the robots to approach the target. For example, the group navigation may be based on a wall-following strategy that, given the structure of the world, guarantees the group to reach the target. However, the fact that the average distance of the team's centroid to the light in the unsuccessful trials is around 20 cm seems to indicate that the absence of walls does not hinder the capability of the group to perform phototaxis. Rather, the walls seem to play an important role during the last phase of the group navigation, when the centroid of the team is required to get into the proximity of the target.

In *Test D*, the group is located in two different types of environment as depicted in figure 6.12. These environments differ from those experienced by the team during the evolutionary phase because the walls are replaced by round obstacles of the diameter of a robot (see figure 6.12). Note that the distance between any two round obstacles allows a robot to pass between them. This means that the probability of robot R_{AL} to collide against one of those round objects depends very much on the spatial configuration of the group during navigation. The results of this test tell us more about the obstacle avoidance capability of the group. The fact that the percentage of unsuccessful trials is around 31% when the light is on the left and only around 11% when the light is on the right indicates that the navigation strategy of the group is quite robust to cope with the kind of environmental changes we made.

6.5 Discussion

The results illustrated in section 6.4 have shown that dynamical neural networks, shaped by artificial evolution, can be successfully used to design homogeneous control structures for a group of morphologically heterogeneous cooperating and communicating robots. Post-evaluation analyses unveiled the mechanisms which underpin the cooperation and coordination of actions of the group. In particular, we focused on the study of the evolved acoustic communication protocol of the best evolved successful group. First, we showed that: (i) all the robots emit sound at a very high intensity; (ii) signalling behaviour is not characterised by oscillatory phenomena; (iii) periodic phenomena, generated by the receiver through a rotational movement associated to the phototaxis, characterise the perception of sound. Then, we proved that oscillations of perceived sound and Interaural Intensity Differences (IIDs) are cues exploited by the robots to generate adaptive actions to safely navigate (i.e., without collisions) towards the target. In particular, the robots exploit these cues to regulate their individual actions with respect to the relative position of sound source(s).

It is reasonable to consider that the evolved behavioural and communication strategies illustrated in section 6.4 are limited to the peculiarities of our simulations. However, our successful results point to the relevance, for the robotics community, of features of our methodological approach which are of more general applicability and re-usable in future research works dealing with the design of homogeneous controllers for groups of heterogeneous cooperating and communicating robots. In particular, we draw the attention of the reader on the following distinctive features of our work: (i) the model of the sound; (ii) the way in which the controllers are wired-up with the sensory apparatus of the robots; (iii) the dynamic speciation of the homogeneous controller, whose mechanisms underpin sensory-motor coordination and social interactions in structurally different agents.

As far as it concerns the model of sound, although inspired by the work of Di Paolo (2000), it presents peculiarities of particular interest. As in (Di Paolo, 2000), and contrary to other experimental works, we did not make use of directional microphones, or any other form of hard-wired/hand-coded mechanisms to discriminate between different sound sources or between self and non-self produced sound. For example, in the work of Marocco and Nolfi (2006b), four directional microphones capture the sound of the nearest robot located within ±45° left or right of each microphone. These types of models not only are hardly portable on a physical system, but they also preclude the possibility to investigate the principles underlying behavioural coordination through sound signalling in a team of autonomous agents. This follows from the fact that in these models, the problem of synchronisation or turn-taking to avoid mutual interference, and of spatial discrimination of sound sources, are eluded thanks to the implementation details. With respect to what described in (Di Paolo, 2000), we strongly simplified the characteristics of the robot's controller. In particular, we did not implement the neural structures which provide the agents in Di Paolo's work the means to further regulate the intensity of emission of sound (i.e., regulation for sound effector, see Di Paolo, 2000) and the receptiveness of the sound sensory neurons (i.e., sensory gain regulation, see Di Paolo, 2000). We simplified the mechanisms to constrain the production of sound by fixing a limit to the intensity of the signal which also corresponds to the saturation level of the sound sensors. That is, the self produced sound can completely saturate the sound sensors of the emitter. Although arbitrarily implemented by the experimenter, these simplifications were introduced to compensate for an increase in structural complexity of the controller due to the nature of the agents' sensory apparatus. In particular, while in the work described in (Di Paolo, 2000) the agents are equipped only with sound sensors, in this work the agents are equipped with sound receptors as well as light or infrared sensors. Moreover, we investigated teams of three robots instead of two robots. Possibly due to these differences, the evolved solutions in Di Paolo's work and in ours diverge significantly. While in Di Paolo's model oscillations and synchronisation in sound production underpin behavioural coordination, in our model, there is no oscillation in sound production.

From an engineering point of view, it is worth to mention that, although extremely effective in terms of collisions, the best evolved navigation strategies are not characterised by a fast phototactic movement. In fact, the strong rotational movement allows for behavioural coordination through sound signalling, while it slows down the movement towards the light. We believe that alternative navigation strategies can potentially be achieved by reintroducing some of the mechanisms originally proposed in (Di Paolo, 2000). These mechanisms facilitate the evolution of oscillatory behaviour in sound production and the distinction between self and non-self components, without having to model phenomena such as time varying frequencies, Doppler effect, etc. A group of robots in which each agent is capable of differentiating between self and non-self and of associating the intensity of the sound perceived in each ear with the distance to the sound source, might employ linear rather than rotational movements. Other hardware specifications, such as the position of the microphones on the robot body, could facilitate the evolution of faster phototactic movement. These issues will be the subject of future investigations.

As far as it concerns the way in which the controllers are wired-up with the sensory apparatus of the robots, we would like to provide further justifications for our implementation choices. Our goal was to generate through artificial evolution a controller capable of guiding both types of robots. For this reason, we chose to keep the group homogeneous with respect to the controllers. That is, at time 0 of each trial, each robot is equipped with exactly the same control structure. However, the properties of the controllers allow for a dynamic speciation: that is, a differentiation of the functionalities of each controller. This differentiation is determined by the attainment, by each controller, of different stable oscillatory dynamics due to the oscillatory pattern experienced through the robot's sensory apparatus. The fact that the CTRNN controller can take time into account to produce the robot's actions (robots can display time-based decision-making) also plays an important and central role in the differentiation of the controllers according to the morphological particularities of each type of robot. Moreover, we wanted to reduce at a minimum the number of parameters which define the search space of the evolutionary algorithm. For this reason, we decided to use neural structures in which the same input neurons in different networks are linked to different type of sensors (see section 6.3.3 for details). Our results suggest that implementation details make possible to generate through artificial evolution homogeneous controllers that can efficiently guide morphologically identical as well as morphologically different groups of robots. In our case, the differences in the flow of sensation coming from different sensory channels (i.e., infrared sensors, ambient light and sound sensors) contribute to induce the specialisation of the controllers with respect to the physical characteristics of the robots, and to the relative role that they play in the group (i.e., the dynamic speciation). This latter mechanism can also be exploited in case of hardware failure, in which an on-line re-assignment of association between agent's sensors and network's input neurons might provide a robust mechanism to preserve the functionality of multi-robot systems. However, in order to efficiently exploit our methodological choices in the latter context, further investigations are required to determine the plasticity of controllers in those circumstances in which they have already undertaken a process of dynamic speciation. That is, it is an open question whether a neuro-controller already specialised to receive as input the reading of a particular set of sensors is capable of redefining its functionalities to guide a robot with a different set of sensors.

6.6 Conclusion

In a context where robots differ in their sensory capabilities, cooperation and coordination of actions of the group are achieved by using an acoustic communication protocol controlled by evolved neural mechanisms. Acoustic signals, determined by the individual emission of a single frequency tone, provide the perceptual cues used by the robots to go beyond the limits of their sensory apparatus in order to obtain robust phototactic strategies as well as obstacle avoidance behaviour. The results of a series of post-evaluation tests carried out on the behavioural strategies of the best evolved group of robots, show interesting operational aspects of the system (see sections 6.4.3, 6.4.4 and 6.4.5). In particular, our analyses highlighted fundamental relationships between the motion of the agents and the appearance of waveforms in sound perception (i.e., affordances, see Gibson, 1977), which are exploited by the robots to mutually coordinate their actions. We also provided evidence that the agents' motion is guided by mechanisms that exploit Interaural Intensity Differences (IIDs): that is, cues used by natural organisms to localise sound sources.

To conclude, from the results of this research work we learn something about how evolution exploits the physics of our system to develop group navigational strategies based on the mutual coordination of actions and cooperation among the agents. The results of this research work are a "proof-of-concept": they show that dynamic artificial neural networks can be successfully synthesised by artificial evolution to design the neural mechanisms required to underpin the behavioural strategies and adaptive communication capabilities demanded by this task. In particular, we developed a sound signalling system that allows a group of morphologically heterogeneous agents that differ in their sensory capabilities to coordinate their actions in order to approach a light bulb without collisions. Post-evaluation analyses unveil operational aspects of the best evolved behaviour. For example, we show that adaptive group behaviour can be achieved without the need of (i) individual built-in mechanisms for distinguishing between self and non-self produced signal, and of (ii) complex neural structures that regulate the turn-taking during communication.

The analysis of evolved individual and social skills give us an estimation of the potentiality of our implementation choices at various levels, from the model of sound to the characteristics of the robots' controller. Although the evolved behavioural and communication strategies may be limited to the peculiarities of this case study, our methodological approach is of more general applicability. In particular, the dynamic speciation of the robots' controllers as well as the elements of the models which bring forth the causal relationships between the physics of the system and the nature of the best evolved collective strategies are contributions of our work that roboticists can employ for the design of more complex forms of social interactions and communication in groups of autonomous robots.

As a closing note, we would like to point out that we are considering alternative tasks and experimental setups involving morphologically heterogeneous robots, which could allow the use of a model of sound which will be closer to the real hardware available (i.e., the *s-bot*) and which will be simple and straightforward to implement. This way, we will be able to validate controllers evolved in simulation in the real world, an endeavour that is always challenging but worth the try since it is the only reliable way to ascertain the functionality and robustness of controllers for autonomous robots.

Chapter 7

Future work and initial experimentation

In this chapter, we present two basic axes of future research that are based on the ideas and experiments detailed in previous chapters.

In section 7.1, we report on initial experimentation aimed at obtaining more complex evolved signals in communicating robots; in particular, we detail the results of an experiment about categorisation and communication. Similar to what was presented in chapter 4, we use principles to obtain autonomous decision-making in a group of two agents that have to collectively categorise their environment. The agents will be again equipped with "ears" and a "mouth" for acoustic communication. However, differently to what was presented in chapter 4, the agents have to categorise each their own sub-part of the environment, to which the other agent has no sensory access; subsequently, the collected information has to be combined and a collective action should be taken.

In section 7.2, we report on initial experimentation aimed to lead towards the evolution of functional self-assembly with autonomous decision-making principles. More specifically, our goal is to design robot controllers that allow robots to connect and disconnect from each other in response to the environmental contingencies. The robots should interact with the rest of the group and with the environment in order to "realise" whether assembly is necessary. The research work that will be presented is an adaptation of the ideas presented and the task considered in chapter 4; moreover, it is strongly correlated with the experiments on self-assembly detailed in chapter 5.

7.1 Towards the evolution of more complex signals

In chapter 4 we presented original work that focused on the interplay between categorisation through integration over time of perceptual experiences and an evolved communication protocol. Communication signals were the result of categories formed on behalf of the robots that are not directly available as sensor inputs. These signals can be called "deictic",[1] as the robots that emit and receive them are sharing the same environment and the signals refer to an environmental state currently available to the robots.

The dance of the honey bee is a striking example from biology where an animal is referring to states of the environment that are not currently available neither to the emitter (dancing bee) nor to the receiver (bees in the nest) (Von Frisch, 1967). This is definitely an example of more complex signalling mechanisms, that humans (and bees) possess, but that do not characterise signalling systems of primates.

> Man is apparently almost unique in being able to talk about things that are remote in space or time (or both) from where the talking goes on. This feature—"displacement"—seems to be definitely lacking in the vocal signalling of man's closest relatives, though it does occur in bee-dancing.
>
> (Hockett, 1960)

In order to obtain more complex signalling capabilities that are closely linked to complex cognitive phenomena as memory, we need to carry out research that aims at this direction. In the following, we will present an experiment whose goal is to explore the evolution of signalling behaviour and communication in a scenario where two robots do not share the same environment but are exploring a sub-part of the world that is not accessible to the other agent.

Even if "displaced" signals, that is, signals that are referring to a context independent of the current sensory experience of the agents (as, e.g., in human language *"I have seen black floor"*) have not been produced, this work might constitute the first step towards understanding how we may arrive to the evolution of communication between agents that do not share the same environment and that can potentially "talk" about their past experiences.

More specifically, this research work is about the evolution of acoustic communication in a two-robot system, in which the agents are required to coordinate their efforts to perform a common task. The aim of this work is to demonstrate the effectiveness of a very simple sound signalling system in a context in which the robots are demanded to share individual experiences to build a common perspective of their world. The robots can communicate by using the global binary signalling system also used in the experiment presented in chapter 4.

We exploit the concept of autonomous decision-making in evolutionary robotics; the designer is not required to make strong assumptions about the essential features on which social interactions are based—e.g., assumptions concerning what communication is and about the requirement of individual competences in the domain of categorisation and naming. The results of the evolutionary process (i.e., the behaviour of the robots and the underlying mechanisms) inform the designer on the effects that the physical interactions among embodied agents and their world have on the evolution of individual behaviour and social skills.

---

[1] According to Norwegian psycholinguist Ragnar Rommetveit, deictic words are *"words that introduce particulars of the speaker's and hearer's shared cognitive field into the message"*.

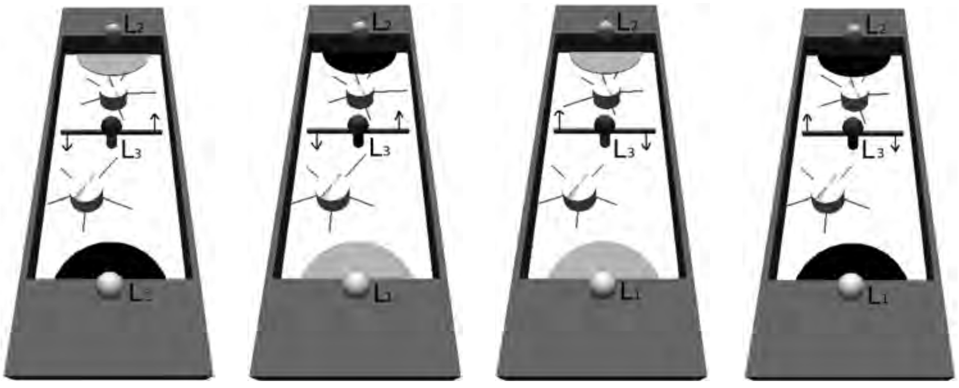
Similarly to (Di Paolo, 2000; Wischmann and Pasemann, 2006), communication is based on the emission in time of asynchronous and mutually determined single tone signals. Differently, in our system the communication concerns events that are not directly available to the sensorial input of the robots at the time it takes place. The results of this work should be taken as a "proof-of-concept" concerning the potentiality of the proposed approach to the design of acoustic communication mechanisms in multi-robot systems. We demonstrate that it is possible to use evolution to define the mechanisms underlying a bi-directional communication protocol based on a very simple acoustic system.

7.1.1 Methods

In the following we present the task in section 7.1.1.1, the simulation model used in section 7.1.1.2, the controller and the evolutionary algorithm in section 7.1.1.3 and the fitness function in section 7.1.1.4.

7.1.1.1 The task

The robot environment is a rectangular arena (120 cm by 50 cm) divided into two equal sides (i.e., upper and lower side) by a horizontal bar that revolves (a revolving door). There are three lights L_1, L_2 and L_3. When L_1 and L_2 are turned on, L_3 is turned off and vice versa. L_1 can only be seen by a robot located in the lower side of the arena while L_2 can only be seen by a robot located in the upper side of the arena. L_3 can be seen from anywhere in the environment. The arena floor is white except in the proximity of L_1 and L_2 up to a distance of 15 cm from the lights, where the floor is painted black or gray. The robots can experience four different combinations of black and gray zones (see figure 7.1). The type of environment in which the robots are located is labelled



E_{10} E_{01} E_{00} E_{11}

Figure 7.1: The four environments E_{10}, E_{01}, E_{00}, and E_{11}. L_1, L_2 and L_3 refer to the lights. The revolving door is indicated by the horizontal bar in the centre of the arena. In each environment, the arrows indicate the direction in which the door revolves. The cylinders with spikes on the white floor represent the robots.

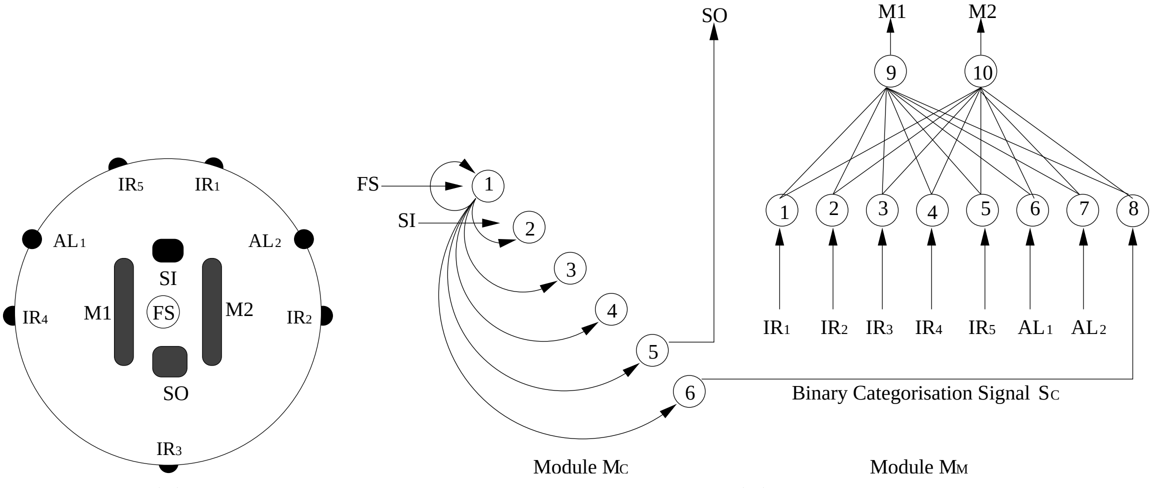
according to the combination of the colour of the floor in the two painted zones. In detail, the environments are labelled E_{ij}, where i corresponds to the colour of the floor in the proximity of L_1 and the second digit to the colour of the floor near L_2. Gray colour corresponds to 0, while black colour corresponds to 1. The four types of environment are: E_{10}, E_{01}, E_{00}, and E_{11}. The revolving door rotates from the horizontal to the vertical position if simultaneously pushed by both robots in the proper direction. Pushing forces exerted by a single robot on the revolving door are not enough to open it. The direction of rotation changes according to the type of environment. The robots have to exert forces to make the door rotate (a) clockwise, if they are located in E_{00} or in E_{11}; (b) anti-clockwise, if located in E_{10} or in E_{10} (see the arrows in figure 7.1).

At the beginning of the first trial and in those that follow an unsuccessful one, the robots are randomly placed in the proximity of L_3. In trials following a successful one, the robots are not repositioned. The sequence of desired actions that each robot is demanded to carry out during a trial can be decomposed into two phases. At the beginning of the first phase, L_1 and L_2 are turned on, the revolving door is in the horizontal position and the colour of the floor in the proximity of L_1 and L_2 is set according to the type of environment that characterises the trial. During this phase, the robots are required to find the painted zone in their side of the white arena floor and remain for at least 6 sec on the painted zone. This exploration is facilitated by the presence of the lights that can be used as beacons (i.e., L_1 for the robot located in the lower side and L_2 for the robot located in the upper side of the arena). The first phase terminates once the 6 sec on the painted zones are elapsed for both robots. At this point, L_1 and L_2 are turned off, L_3 turned on, and the second phase begins. In the second phase, the two robots are required to move back towards the middle of the arena, approach the revolving door, and simultaneously push the door in order to open it and to reach the previously inaccessible opposite side of the arena. As mentioned above, the direction of rotation changes according to the type of environment. Therefore, in order to rotate the revolving door from the horizontal towards the vertical position, the robots are required to exchange information about the colour of the floor in the proximity of the light—L_1 or L_2—previously approached. A trial successfully terminates once both robots, by rotating the revolving door, move into the opposite side of the arena, and reach a distance of 24 cm from L_3. At the end of a successful trial, L_3 is turned off, L_1 and L_2 are turned on, the rotating door automatically returns to the horizontal position and a new trial begins. A trial is considered unsuccessful if a single robot exerts forces in both arms of the revolving door (i.e., west and east of L_3). This behaviour, referred to as trial-and-error strategy, is penalised by the fitness function (see section 7.1.1.4).

Note that this task requires coordination of actions, cooperation and communication between the robots in order to open the revolving door. For each robot, the perception of a gray or black floor can be associated both to a clockwise and anti-clockwise rotation of the revolving door. Only the combination of the two coloured zones unambiguously identifies a rotational movement. Since a robot can only walk on a single zone per trial, the task can be successfully accomplished in all the environmental conditions only by a group of robots that communicate through sound. Without communication, a single robot can only exploit a trial-and-error strategy. By using a simple sound signalling system the robots should inform each other on the colour of the floor in the proximity of the light they perceive—L_1 or L_2—and consequently push the door in the proper direction as explained above. Finally, another important point is that the action of the robots with respect to correctly positioning themselves at the side of the revolving door, is not a reaction to the acoustic communication. Rather, the robots after the communication has taken place, have to retain the information they have gathered about the environment until they arrive to the vicinity of the revolving door and then assume the appropriate positions. In other words, the communication act provides information that has to be kept in the memory of the agents until they arrive to the point where they can pursue the correct action, with respect to the environmental contingencies.

7.1.1.2 The simulation model

The robot and its world are simulated using simulation software based on the Open Dynamic Engine (see `http://www.ode.org/`), a 3D rigid body dynamics simulator that provides primitives for the implementation of detailed and realistic physics-based simulations. This choice was driven by the need to simulate the dynamics of the revolving door and the effect of the pushing forces robot exert on it. Our simulation models some of the hardware characteristics of the real *s-bots* (see Mondada et al., 2004, and section 2.3.1, for details). Our simulated robot has a differential drive motion provided by a traction system composed of four wheels: two lateral, motorised wheels and two spherical, passive wheels placed in the front and in the back, which serve as support. The four wheels are fixed to



(a) (b)

Figure 7.2: (a) The simulated robot. $IR_i, i \in \{1, 2, \ldots, 5\}$ are the infrared sensors; $AL_i, i \in \{1, 2\}$ are the ambient light sensors; FS is the floor sensor; SI is the sound sensor (i.e., the microphone); SO is the sound actuator (i.e., the loudspeaker); M_1 and M_2 are respectively the left and right motor. (b) The network architecture: module M_C and module M_M. For M_C only the efferent connections for one neuron are drawn. S_C is the binary categorisation signal sent, at each updating cycle, by M_C to M_M.

the cylindrical body that holds the sensors. In particular, robots make use of 5 infrared sensors IR_i, two ambient light sensors AL_i, a floor sensor FS, a loudspeaker SO to emit sound and an omni-directional sound sensor SI to perceive sound (see figure 7.2a). Light levels change as a function of the robot's distance from the lamp. FS, placed underneath the robot, detects the level of gray of the floor. It outputs the following values: 0 if the robot is positioned over white floor; 0.5 if the robot is positioned over gray floor; 1 if the robot is positioned over black floor. SO produces a binary output (on/off). SI has no directionality and intensity features. 10% uniform noise is added to IR_i and AL_i readings, the motor outputs and the position of the robot.

7.1.1.3 The controller and the evolutionary algorithm

The controller of each agent is composed of two modules referred to as M_C and M_M (see figure 7.2b). The modularisation is hand-coded to facilitate the evolution of successful behavioural strategies. M_C is a non-reactive module, that is, a six neurons fully connected CTRNN (see chapter 3, section 3.1.2 and Beer and Gallagher, 1992). M_C should detect in which type of environment the robot is currently located. The categorisation has to be based on the FS's readings of both robots. Thus, it demands communication between the agents. For this reason, M_C takes input from FS and SI and it outputs the state of the SO and S_C (i.e., the binary categorisation signal). In other words, at every updating cycle, M_C is in charge of (a) managing sound by producing the signal the robot emits and by receiving the signal of either robot, and (b) "informing" M_M on the type of environment in which the robot is currently located by setting the value of the binary categorisation signal S_C either to 0 or 1. M_M is a reactive module, that is, a feed-forward artificial neural network made of eight sensory neurons and two output neurons. M_M is expected to (a) guide the robot avoiding collisions with the arena walls, and (b) "parse" the value of S_C to determine in which side to push the revolving door (i.e., anti-clockwise if current trial in E_{10} or E_{01}, clockwise if current trial in E_{00} or E_{11}, see also figure 7.1). M_M takes input from IR_i, $i \in \{1, 2, \ldots, 5\}$, from AL_i, $i \in \{1, 2\}$, and S_C, and it outputs the speed of the robot's wheels. The following associations (a) $S_C = 1$, robots located in E_{10} or E_{01}, anti-clockwise rotational direction of the revolving door, and (b) $S_C = 0$, robots located in E_{00} or E_{11}, clockwise rotational direction of the revolving door, are determined a priori by the experimenter (see section 7.1.1.4). The neural mechanisms and the communication protocol required by the robots to build these relationships from the sensors' readings are set by evolution.

The states of the neurons of M_C and M_M are governed by the following equations, respectively:

$$\frac{dy_i}{dt} = \frac{1}{\tau_i}\left(-y_i + \sum_{j=1}^{6} \omega_{ji}\sigma(y_j + \beta_j) + gI_i\right), \; i \in \{1, 2, \ldots, 6\}; \; \sigma(x) = \frac{1}{1 + e^{-x}} \tag{7.1}$$

$$dy_i = \begin{cases} -y_i + gI_i & i \in \{1, 2, \ldots, 8\} \\ \left(-y_i + \sum_{j=1}^{8} \omega_{ji}\sigma(y_j + \beta)\right) & i \in \{9, 10\}; \end{cases} \tag{7.2}$$

where, using terms derived from an analogy with real neurons, y_i represents the cell potential, τ_i is the decay constant, g is a gain factor, I_i the intensity of the sensory perturbation on sensory neuron i, ω_{ji} the strength of the synaptic connection from neuron j to neuron i, β the bias term, $\sigma(y_j + \beta)$ the firing rate. The parameters ω_{ji}, τ, β and g are genetically encoded. Cell potentials are set to 0 any time the network is initialised or reset, and circuits are integrated using the forward Euler method with an integration step-size of 0.1. Note that the cell potentials of M_M's neurons do not depend on time. That is, the neurons' decay constant τ is set to 0.1, as the integration step-size dt. In M_C, the cell potentials y_i of the 5^{th} and the 6^{th} neuron, mapped into [0,1] by a sigmoid function σ, set the state of the robot's sound actuator SO and of the binary categorisation signal S_C. The robot emits a sound if $SO \geq 0.5$. $S_C = 1$ if $\sigma(y_6 + \beta_6) \geq 0.5$ otherwise $S_C = 0$. In M_M, the cell potentials y_i of the 9^{th} and the 10^{th} neuron, mapped into [0,1] by a sigmoid function σ and then linearly scaled into $[-6.5, 6.5]$, set the robot motors output.

A simple generational genetic algorithm is employed to set the parameters of the networks (Goldberg, 1989). The population contains 80 genotypes. Generations following the first one are produced by a combination of selection with elitism, recombination and mutation. For each new generation, the three highest scoring individuals ("the elite") from the previous generation are retained unchanged. The remainder of the new population is generated by fitness-proportional selection (also known as roulette wheel selection) from the 64 best individuals of the old population. Each genotype is a vector comprising 67 real values, chosen uniformly random from the range $[0, 1]$. The first 18 genes are used to set the parameters of M_M (i.e., 16 connection weights, 1 bias term and 1 gain factor both shared by all the input neurons). The other 49 genes are used to set the parameters of M_c (i.e., 36 connection weights, 6 decay constants, 6 bias terms, and 1 gain factor).

7.1.1.4 The fitness function

During evolution, each genotype is translated into a robot controller (i.e., modules M_C and M_M, see section 7.1.1.3), and cloned in each agent. Then, the two robot group is evaluated two times in each environment type E_{11}, E_{00}, E_{01}, and E_{10}, for a total of eight trials. Note that the sequence order of the environment type experienced by the robots—randomly chosen at the beginning of each generation—has a bearing on the overall performance of the group since the robots' controllers are reset only at the beginning of the first trial. Each trial differs from the others in the initialisation of the random number generator, which influences the robots' starting position and orientation anytime the robots are initialised, and the noise added to motors and sensors. The robots are randomly placed in the arena at the beginning of the first trial and repositioned in subsequent trials following an unsuccessful one. Within a trial, the robots life-span is 90 simulated seconds (900 simulation cycles). A trial is terminated earlier in case a robot crashes with the arena walls, or if the group successfully accomplishes its task. For each trial $e \in \{1, 2, \ldots, 8\}$, the group is rewarded by an evaluation function which seeks to assess the ability of the robots to open the revolving door located at the centre of the arena (see section 7.1.1.1). This requires the robots to be able to determine the nature of the environment (i.e., E_{11}, E_{00}, E_{01}, or E_{10}) by using acoustic communication. The final fitness F attributed to a group controlled by a specific genotype is the average group score over a set of eight trials, and it is computed as follows:

$$F = \left(\frac{\sum\limits_{e=1}^{8} FM_e}{8}\right) + FC; \quad FM_e = \frac{\sum\limits_{r=1}^{2}(\psi_r + fm_r\kappa_r\rho_r)}{2}, \; FM_e \in [0, 2.4]; \tag{7.3}$$

$$FC = \begin{cases} 0, & \text{if } \frac{\left(\sum\limits_{e=1}^{8}\sum\limits_{r=1}^{2} fc_{er}\right)}{2} \leq 0; \\ \frac{\sum\limits_{e=1}^{8}\sum\limits_{r=1}^{2} fc_{er}}{2}, & \text{otherwise;} \end{cases}, \; FC \in [0, 1.0] \tag{7.4}$$

$\psi_r = 0$ if robot r did not terminate the first phase of a trial, otherwise $\psi_r = 1$. $\rho_r = \frac{1}{5}$ if robot r collided with the arena walls, otherwise $\rho_r = 1$. $fm_r = 1.0 - (d_{rL_i})$ with d_{rL_i} corresponding to the normalised distance between the robot r and the light L_i. During the first phase of a trial $i = 1$ for the robot r located in the lower side of the arena, and $i = 2$ for the robot r located in the upper side of the arena. During the second phase of a trial, $i = 1$ for the robot r located in the upper side of the arena, and $i = 2$ for the robot r located in the lower side of the arena. $\kappa_r = 1$ if robot r didn't terminate the first phase of a trial, or if, after having done so, it exerts pushing forces in the rotational direction of the revolving door (see the arrows in figure 7.1). $\kappa_r = 0.5$ if robot r, after having terminated the first phase of a trial, it exerts pushing forces in a direction opposite to the rotational direction of the revolving door. $fc_{er} = \frac{\sum_{s=(t_c)}^{T}(P_{s-tc})}{T-t_c}$, where t_c corresponds to the simulation cycles at 10 seconds after the end of the first part of the task, T corresponds to the simulation cycles at the end of the trial e and

$$P_{s-tc} = \begin{cases} +1 \text{ if } ((E_{11} \vee E_{00}) \wedge (S_C = 0)) \vee ((E_{10} \vee E_{01}) \wedge (S_C = 1)) \\ -1 \text{ if } ((E_{11} \vee E_{00}) \wedge (S_C = 1)) \vee ((E_{10} \vee E_{01}) \wedge (S_C = 0)) \end{cases} \tag{7.5}$$

In other words, during the first phase of a trial, FM_e rewards the robots for approaching the light at the corresponding side of the arena (i.e., L_1 for robot in the lower side; L_2 for robot in the upper side). During the second phase of a trial, FM_e rewards the robots for approaching the opposite side of the arena. FC rewards the robots for setting the state of the fifth neuron (i.e., y_5) of M_C so that (a) S_C results equal to 0 during the second phase of trials in E_{00} and E_{11}; (b) S_C results equal to 1 during the second phase of trials in E_{10} and E_{01} (see also section 7.1.1.3). Note that F does not refer anyhow to signalling behaviour. F rewards the robots for accomplishing the task as detailed in section 7.1.1.1. However, due to the nature of the task, the robots can be successful only if they coordinate their actions using the sound signalling system. By leaving signalling behaviour out of the fitness function, we clean our model from preconceptions concerning what (i.e., semantics) and how (i.e., syntax) successful groups communicate, and we let evolution determine the characteristics of the communication protocol.

7.1.2 Results

Ten evolutionary simulations, each using a different random initialisation, were run for 4.800 generations. Given the nature of the fitness function, the highest fitness score that a group can reach is 3.4. This score corresponds to the behaviour of a group in which each robot (i) finds the coloured zone on the white arena floor; (ii) communicates to the robot at the opposite side of the arena the colour encountered in its side; (iii) uses the combination of colours to properly set the binary categorisation signal S_C; and (iv) pushes the revolving door in the proper direction until it reaches the opposite side of the arena.

Notice, however, that fitness scores lower than 3.4 might be associated to equally successful alternative strategies. In particular, successful strategies may have the fitness component $FC < 1$ in case M_C doesn't correctly set S_C for the entire length of the time interval from t_c to T as demanded by the fitness function (see section 7.1.1.4). For a group to be successful, what matters is that (i) M_C is capable of discriminating environments in which the door revolves clockwise from those in which the door revolves anti-clockwise; (ii) this discrimination is made available to M_M through the value of S_C; (iii) differences in time of S_C's reading induce different behavioural responses. How these processes are implemented may vary with respect to the nature of the mechanisms found by evolution. Not all the implementations which allow a group to be successful get the highest fitness score.

Thus, in order to have a better estimate of the behavioural capabilities of the best evolved controllers, we post-evaluate, for each run, the genotype with the highest fitness. These groups are referred to as $g_i, i \in \{1, 2, \dots, 10\}$. The entire set of post-evaluations (i.e., 2400 trials, 100 evaluations for each permutation, 100*N! with N=4) should establish whether a group of robots is capable of accomplishing the task as described in section 7.1.1.1 in all four types of environment.

The results of the post-evaluation tests are shown in table 7.1. The data show that only two groups (g_2 and g_4) have a very high success rate (higher than 98%) in all four types of environment (see figure 7.1b, gray rows); g_1, g_3, g_5, g_8 and g_9 are capable of carrying out the task only when the door revolves clockwise, and g_{10} only when the door revolves anti-clockwise; g_6 and g_7 fail in only one type of environment.

From a behavioural point of view, the failures are due to trial-and-error strategies. That is, during the second phase of the task, both robots push the revolving door both west and east of L_3 instead of exerting forces directly on the proper side of the bar (see table 7.2).

From table 7.3, we can deduce that failures due to collisions are very rare. The lower success rate of g_{10} in E_{00} and E_{11} is mainly due to the fact that the robots of this group are not able to exert enough forces to rotate the revolving door. From a mechanism point of view, the failure of each single robot can be caused by either (a) M_C failing to correctly categorise the environment by properly setting S_C as made explicit in section 7.1.1.3 or (b) M_M failing to "interprete" the value of S_C as produced by M_C. Post-evaluation tests show that for almost all the unsuccessful groups, it is M_C that by setting incorrectly the value of S_C, does not allow M_M to choose the correct direction of rotation of the revolving door (see table 7.4). It seems that robots of unsuccessful groups are not capable of informing each

	(%) Success			
	E_{10}	E_{01}	E_{00}	E_{11}
g_1	0.0	0.0	100.0	100.0
g_2	99.6	99.8	100.0	98.16
g_3	0.0	0.0	97.8	96.3
g_4	100.0	100.0	100.0	99.6
g_5	0.0	0.0	79.7	82.5
g_6	99.0	94.2	92.7	0.0
g_7	99.5	99.5	100.0	0.0
g_8	100.0	100.0	0.0	0.0
g_9	0.0	0.0	100.0	99.8
g_{10}	100.0	99.7	0.0	0.0

Table 7.1: Results of post-evaluation tests, showing for the best evolved groups of each run the percentage of successful trials in each type of environment. In gray the successful groups.

Table 7.2: Results of post-evaluation tests, showing for the best evolved groups of each run the percentage of unsuccessful trials per robot in each type of environment due to trial-and-error strategy. This corresponds to a robot exerting forces in both arms of the revolving door (i.e., west and east of L_3) instead of touching the bar on the correct side. In gray the successful groups.

	(%) Failure due to trail-and-error strategy							
	E_{10}		E_{01}		E_{00}		E_{11}	
	R_1	R_2	R_1	R_2	R_1	R_2	R_1	R_2
g_1	100.0	100.0	100.0	100.0	0.0	0.0	0.0	0.0
g_2	0.0	0.0	0.0	0.0	0.0	0.0	1.8	1.8
g_3	99.3	99.33	99.5	99.5	0.0	0.0	0.0	0.0
g_4	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.3
g_5	99.7	99.7	99.5	99.5	1.3	1.7	0.0	0.0
g_6	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0
g_7	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0
g_8	0.0	0.0	0.0	0.0	100.0	100.0	100.0	100.0
g_9	100.0	100.0	100.0	100.0	0.0	0.0	0.0	0.0
g_{10}	0.0	0.0	0.0	0.0	32.17	44.7	100.0	100.0

other about the colour of the painted zone in the proximity of L_1 and L_2. Consequently, in the absence of an effective communication protocol, it turns out impossible for M_C to properly set S_C. In the following paragraphs, we analyse the communication protocol used by a successful group.

Figure 7.3a illustrates the structures of signalling behaviour of the successful group g_4. In this post-evaluation test, the group undergoes 4 trials with the environment presented

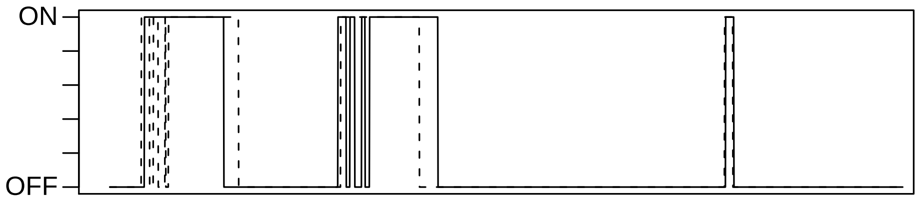
Table 7.3: Results of post-evaluation tests, showing for the best evolved groups of each run the percentage of unsuccessful trials per robot in each type of environment due to collisions. In gray the successful groups. As we can see from the table, failures due to collisions are very rare.

	(%) Collisions							
	E_{10}		E_{01}		E_{00}		E_{11}	
	R_1	R_2	R_1	R_2	R_1	R_2	R_1	R_2
g_1	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.17
g_2	0.17	0.00	0.17	0.0	0.00	0.00	0.00	0.00
g_3	0.17	0.50	0.00	0.5	0.00	0.33	0.00	1.00
g_4	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00
g_5	1.50	0.33	0.33	1.5	0.33	0.50	0.17	0.17
g_6	0.17	0.00	0.00	0.0	0.00	0.17	0.00	0.00
g_7	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00
g_8	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00
g_9	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.17
g_{10}	0.00	0.00	0.00	0.0	0.00	0.00	0.00	0.00

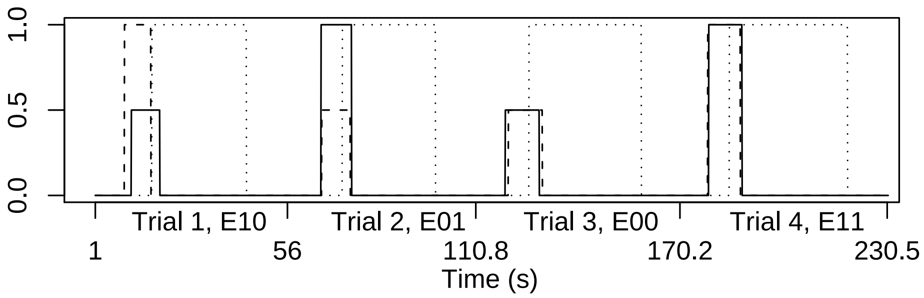
Table 7.4: Post-evaluation analysis of the 10 best groups. In gray the successful groups. This table shows the percentage of the correctness of the binary categorisation signal S_C which is very high ($> 95\%$) for all environments and both robots only for the two successful evolutionary runs g_2 and g_4.

	(%) of correctness of signal S_C							
	E_{10}		E_{01}		E_{00}		E_{11}	
	R_1	R_2	R_1	R_2	R_1	R_2	R_1	R_2
g_1	0.00	0.01	0.00	0.00	99.99	99.98	99.92	99.88
g_2	99.64	95.76	95.70	99.71	100.00	100.00	98.26	98.27
g_3	0.72	0.71	0.59	0.00	99.56	99.45	99.02	98.99
g_4	100.00	100.00	100.00	100.00	100.00	100.00	99.72	99.74
g_5	5.12	4.37	3.53	6.16	93.89	93.05	95.22	93.20
g_6	99.83	99.83	99.33	99.33	94.64	94.42	0.00	0.00
g_7	99.98	99.88	99.88	100.00	100.00	100.00	0.13	0.48
g_8	100.00	100.00	100.00	100.00	0.00	0.00	0.00	0.00
g_9	0.00	0.00	0.00	0.00	100.00	100.00	99.92	99.92
g_{10}	98.58	90.024	89.86	98.49	100.00	100.00	5.94	5.93

in the following sequence: E_{10}, E_{01}, E_{00}, and E_{11}. In each trial the robots do not emit sound before reaching the coloured zones. The perception of gray does not induce the emission of sound. Therefore, in E_{00} no robots emit sound (see figure 7.3 trial 3). The absence of sound in the environment lets M_C set S_C to 0 in both robots. $S_C = 0$ is correctly "interpreted" by M_M modules so that both robots push the revolving door clockwise. The



(a)



(b)

Figure 7.3: Post-evaluations of group g_4. Dashed lines refers to the robot placed at the beginning of trial 1, in the upper side of the arena; continuous lines refer to the robot placed in the lower side of the arena. (a) Sound signals. (b) Floor sensors readings. Dotted line indicates the state of L_3, 1 = ON, 0 = OFF. On the x axis is indicated the time of start and end of each trial.

perception of a black zone induces the robots to emit intermittent bursts of sound (see figure 7.3a trials 1, 2 and 4). In trials E_{10} and E_{01}, the perception of these intermittent bursts induces the robot that is on gray to emit a continuous tone. The perception of a continuous tone induces the robot on black to imitate its fellow, so that at the time when L_3 turns on (see figure 7.3b, dotted line) both robots emit a continuous tone. The presence of sound in the environment lets M_C set S_C to 1 in both robots. $S_C = 1$ is correctly "interpreted" by M_M modules so that both robots push the revolving door anti-clockwise. Both robots autonomously stop emitting sound before the end of a trial in E_{10} or E_{01}, few seconds after the aperture of the revolving door. Thus, at the beginning of the following trial both robots are in the state of not emitting sound. In trials E_{11}, the asynchronous emission of intermittent bursts of sound by both robots determines moments of silence which inhibit signalling behaviour. At the time when L_3 turns on, none of the robots is signalling. The absence of sound in the environment lets M_C set S_C to 0 in both robots. $S_C = 0$ is correctly "interpreted" by M_M modules so that both robots push the revolving door clockwise as in E_{00}.

7.1.3 Conclusion

We described a model in which artificial evolution is employed to design neural mechanisms that control the motion of autonomous robots required to communicate through sound to perform a common task. The results of this work are a “proof-of-concept”: they demonstrate that evolution can exploit a simple sound system, detailed in section 7.1.1.3, to design the mechanisms that allow two robots cooperate by using bi-directional acoustic interactions. Post-evaluation tests illustrate the nature of the robots’ communication protocol based on entirely evolved asynchronous and mutually determined single tone signals.

Concerning future work, we believe that priority should be given to investigations aimed at limiting the amount of a priori assumptions that we have been forced to make in this first study. In particular, we are referring to the modularisation of the control structures and the arbitrary associations detailed in section 7.1.1.3. This modularisation was introduced in order to obtain solutions to the problems at hand; a fully integrated approach did not lead to satisfactory solutions. This might be linked to bootstrapping problems, related to the fitness function design, or, alternatively, to a difficulty introduced by the fact that the robots have to share their communication medium, using the signalling system also described in section 4.3.2. In future work, we will investigate a scenario where the robots use a different audio channel to emit signals, e.g., by emitting at different frequencies (see chapter 4, section 4.6.3—this modification can allow robots to distinguish between *self* and *non-self* produced sound, and its implementation is feasible on the *s-bot*, for two robots emitting at two distinct frequencies). Such a scenario will significantly simplify the automatic design of the signalling protocols and might facilitate the evolution of integrated controllers. Similarly, the use of sound signalling could be replaced by the use of the omni-directional camera mounted on the turret of the *s-bot*, in a similar way as it was used in section 5.3.2. The use of the camera could replace both the infrared sensors and the sound sensors and actuators, since (i) the camera returns the distance of the objects it perceives and (ii) the robots can light up in different colours (red, green and blue) which can be discriminated by the camera and serve as distinct communication signals.

This work can be seen as a first step in the investigation of the evolution of more complex signals in groups of robots. We did not observe the emergence of displaced signals—signals which convey information about the environment which is independent from the current sensory state of the signalling robot. Instead, robots signal immediately after they perceive the colour of the floor in the proximity of the lights. This bi-directional communication is informing them on the position they should assume at the rotating door. More speculation and further experimentation is needed in order to understand how artificial agents can i) “store” in their memory structures information about the environment they perceive; ii) share this information upon encounter with other agents at a later time.

Finally, we would like to test this asynchronous single tone communication protocols on the real hardware (i.e., the *s-bots*). This is an endeavour which has not been yet pursued, as none of the research works treating this issue have presented real robot experiments (see

Figure 7.4: Chain formation of ants of the species *Œcophilla longinoda*. Image adapted from (Anderson et al., 2002).

Di Paolo, 2000; Wischmann and Pasemann, 2006, for examples). Concerning the nature of our task, which demands many modifications in the environment (turning the lights on/off) and coordinated physical interactions involved with pushing to open the revolving door, it is of course a large constraint that makes the experimentation on real hardware not so straightforward. We are currently speculating on a simpler setup which can allow the evolution of similar communication protocols but which can facilitate the testing on real robots.

7.2 Towards the evolution of functional self-assembly

In chapter 5 we have studied minimal mechanisms underlying self-assembly: how two agents can coordinate their movements in order to decide which robot will grip the other. In the work presented in that chapter, the formation of an assembled structure was the goal. However, in nature (e.g., ants of the species *Œcophilla longinoda*), self-assembly typically takes place in order to accomplish some function. For example, ants build chains composed of their own bodies to bridge gaps (see Lioni et al., 2001; Anderson et al., 2002, and figure 7.4). Similarly, in order for the robots to choose to move collectively as an assembled structure, the environment must present some contingencies that demand it. Trianni et al. (2004) define functional self-assembly as the phenomenon of *"robots physically connecting to each other, any time environmental contingencies prevent a single robot to achieve its goal"*.

In chapter 4 we studied the switch from solitary to social behaviour via communicative acts which result from individual environmental categorisation. The environment where the robots were located did not explicitly demand the use of social behaviour, no matter how this is defined: the task could be solved by the robots either individually or by cooperation. Imagine however an environment where a connected structure could be able to achieve objectives that a single robot could not. Trianni and Dorigo (2005) and Groß et al. (2006a) have demonstrated on real hardware how connected *s-bots* (i.e., a *swarm-bot*) can pass over small troughs, that is, can bridge a gap. Time-dependent decision-making mechanisms as those developed in chapter 4 can help the robots "realise" that social behaviour in the form of self-assembly is required in order to achieve their objectives. In that case of course, the environment would explicitly require cooperation.

In this section we present a first step towards the evolution of functional self-assembly: we extend the research work presented in chapter 4 by coming up with a similar experimental setup and by requiring the robots to exhibit simple collective behaviours if the environment requires them, in the form of aggregation. In the future, we wish to expand this work by demanding that the robots i) self-assemble, ii) move coordinately to achieve the group's objective (for example cross a trough too large for a single robot to cross etc.). This we will do by exploiting knowledge acquired during our experimentation on the evolution of self-assembly, as presented in chapter 5.

An initial investigation of functional self-assembly in the context of collective robotics has been performed by Trianni and Dorigo (2005). The authors performed research in an abstract simulation environment and managed to evolve neuro-controllers able to bring forth assembly and disassembly in a group of robots, according to the state of the part of the environment in which they were located. Notice that the states of the environment were directly available as inputs to the robots and that they were a priori associated to social (group assembled) or solitary (group non-assembled) behaviour by the experimenters.

In the work we present in the following, which will be the basis of on-going experimentation, we demand that the robots themselves determine if assembly is necessary or not. Moreover, our experimental setup is oriented towards obtaining controllers that can be ported on real hardware. More specifically, we will use the simulated environment employed in the experiments on self-assembly detailed in chapter 5.

7.2.1 Methods

In this section we will present the task (see section 7.2.1.1), the simulation model used in section 7.2.1.2, the controller and the evolutionary algorithm in section 7.2.1.3 and the fitness function in section 7.2.1.3.

7.2.1.1 The task

The path towards the evolution of neural controllers for functional self-assembly in a physical *swarm-bot* passes through the definition of the following experimental scenario. A group of *s-bots* is placed in an arena that is surrounded by some obstacles that *s-bots* cannot overcome individually. The arena may have a *way out*, that is, a passage through which a solitary *s-bot* can exit (see figure 7.5a). However, an *s-bot* does not have the perceptual abilities to detect the *way out* from every location in the arena. Therefore, *s-bots* should first search for the *way out* and, if they do not find any as in figure 7.5b, they should aggregate and self-assemble in order to collectively overcome the obstacles that surround the arena. As mentioned above, we consider in this work only the first part of this scenario concerning the decision to switch from the individual behaviour of searching for the *way out* to the collective behaviour of aggregating in one place. The second part of the scenario concerning self-assembly is on-going work.

It should be noted that this task is very similar to the task we tackled in chapter 4. We introduced some modification (*way out* instead of *way in* zone, absence of a light, three robots instead of two, etc.) because we believe that the new setup is more suited to the study of functional self-assembly.

7.2.1.2 The simulation model

The simulator used to evolve the required behaviour relies on a specialised 2D dynamics engine (see Christensen, 2005), and we use a simulated version of the real *s-bot* (see figure 7.6 which we reproduce again for the sake of completeness, and chapter 2.3.1 for details).

Each *s-bot* is provided with four proximity sensors placed under the chassis—referred to as *ground sensors*—that can be used for perceiving the ground's grey level. When the sensor is placed over white ground, it returns a high value due to the high reflectivity of the ground. On the contrary, if the ground colour is black, the reflectivity is low and consequently the sensor returns a value close to 0. The raw sensor readings are recorded and scaled in the interval [0,1] before being processed by the neural controller.

Each robot is also equipped with an omni-directional camera, which is used to perceive the presence and the corresponding distance of neighbouring *s-bots*. The omni-directional

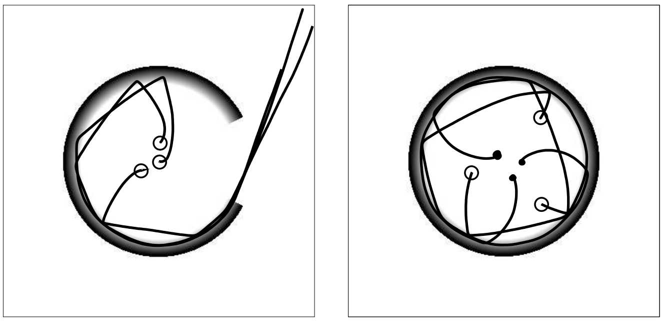

(a) (b)

Figure 7.5: (a,b) The experimental arena contains a circular band in shades of grey, which may or may not have the *way out*. Dark lines represent the trajectories of the three *s-bots*, and empty circles indicate their starting positions.

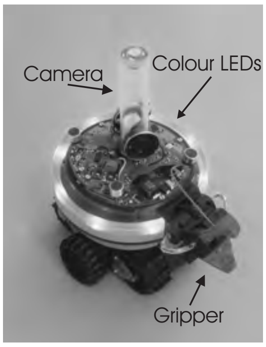


Figure 7.6: Image of an s-bot.

camera can perceive the red colour continuously emitted by the *s-bots* by means of their coloured LEDs embedded in the T-shaped ring (see figure 7.6). The circular image obtained from the camera is filtered in order to extract only the red objects. Then, it is split in 4 sectors of 90° each (front-left, front-right, rear-left, rear-right) and the distance of the closest red object in each sector is computed. With such a system, the closest *s-bot* in each sector can be perceived up to a distance of about 50 *cm*. Also in this case, distances are scaled in the interval $[0, 1]$ before being processed by the neural controller.

In order to communicate with each other, *s-bots* are provided with a very simple signalling system, which can produce a continuous tone with fixed frequency and intensity. When a tone is emitted, it is perceived by every robot in the arena, including the signalling *s-bot*. The tone is perceived in a binary way, that is, either some *s-bot* is signalling in the arena, or no one is, as was done in chapter 4.

Notwithstanding the efforts to devise a precise simulation, some characteristics of the robots and of the robot-environment interaction may escape the modelling phase. For this reason, noise is used to ensure that the evolved behaviour will cope with differences between simulation and reality (Jakobi, 1997). Except for the binary communication system, noise is simulated for all sensors and actuators, adding a random value uniformly distributed in the interval $[-5\%, 5\%]$ with respect to the maximum value.

Three *s-bots* are initially placed up to 25 cm from the centre of a boundless arena. The arena contains a circular band in shades of grey (inner radius: 1.0 m; outer radius: 1.2 m—see figure 7.5). The outer border of the circular band is painted in black and simulates the presence of a trough/obstacle that the *s-bots* cannot overcome individually: the simulation is stopped whenever *s-bots* pass over the black border, and the trial is considered unsuccessful. The grey level of the circular band can be perceived by the *s-bots* only locally through the ground sensors. It is meant to warn *s-bots* about the presence of the simulated trough/obstacle: the darker the ground colour, the closer the danger. The *s-bots* can be placed in two different environments: in *Env.A*, the circular band is discontinuous—i.e., there is a *way out* through which the *s-bots* can exit (see the trajectories in figure 7.5a). In *Env.B*, the *way out* is not present and therefore *s-bots* should aggregate after having searched for it (see the trajectories in figure 7.5b). The amplitude of the *way out* is randomly selected in each trial within the interval $[\pi/4, \pi/2]$.

7.2.1.3 The controller and the evolutionary algorithm

Homogeneous groups of *s-bots* are controlled by artificial neural networks, whose parameters are set by an evolutionary algorithm. A single genotype is used to create a group of individuals with an identical control structure. Each *s-bot* is controlled by a CTRNN (see chapter 3, section 3.1.2 and Beer and Gallagher, 1992). The neural network has a multi-layer topology, as shown in figure 7.7: neurons $N_{I,1}$ to $N_{I,9}$ take input from the robot's sensory apparatus, neurons $N_{O,1}$ to $N_{O,3}$ control the robot's actuators, and neurons $N_{H,1}$ to $N_{H,5}$ form a fully recurrent continuous time hidden layer. The input neurons are simple relay units, while the output neurons are governed by the following equations:

$$o_j = \sigma(O_j + \beta_j), \qquad O_j = \sum_{i=1}^{5} W_O(i,j)\,\sigma(H_i + \beta_i), \qquad \sigma(z) = \frac{1}{1+e^{-z}}, \tag{7.6}$$

where, using terms derived from an analogy with real neurons, O_j and H_i are the cell potentials of respectively output neuron j and hidden neuron i, β_j and β_i are bias terms, $W_O(i,j)$ is the strength of the synaptic connection from hidden neuron i to output neuron j, and o_j and $h_i = \sigma\left(H_i + \beta_i\right)$ are the firing rates. The hidden units are governed by the following equation:

$$\frac{dH_j}{dt} = \frac{1}{\tau_j}\left(-H_j + \sum_{i=1}^{5} W_H(i,j)\sigma(H_i + \beta_i) + \sum_{i=1}^{9} W_I(i,j) I_i\right), \tag{7.7}$$

where τ_j is the decay constant, $W_H(i,j)$ is the strength of the synaptic connection from hidden neuron i to hidden neuron j, $W_I(i,j)$ is the strength of the connection from input neuron i to hidden neuron j, and I_i is the intensity of the sensory perturbation on neuron i.

Four input neurons—$N_{I,1}$ to $N_{I,4}$—are set looking at the four sectors of the image grabbed by the omni-directional camera. Four other input neurons—$N_{I,5}$ to $N_{I,8}$—are set directly from the four ground sensors. Finally, input neuron $N_{I,9}$ is a binary input set by the perception of a sound signal. The neurons $N_{O,1}$ and $N_{O,2}$ are used to set the speed of the *s-bot*'s wheels. Neuron $N_{O,3}$ is used to set the state of the loudspeaker, which is turned on if the neuron output is higher than 0.5, and off otherwise. The weights of the connection between neurons, the bias terms and the decay constants are genetically encoded parameters. Cell potentials are set to 0 each time a network is initialised or reset. State equations are integrated using the forward Euler method with an integration step-size of 0.1 seconds.

In order to set the parameters of the *s-bot*' controllers, a simple generational evolutionary algorithm is employed (Goldberg, 1989). The population contains 100 genotypes that

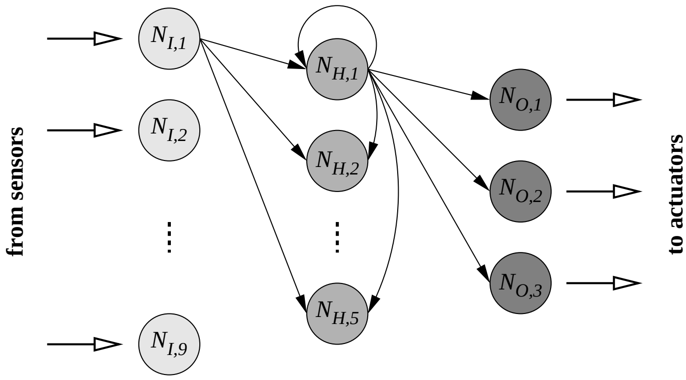


Figure 7.7: The multi-layer topology of the neural neural controller. The hidden layer is composed of continuous time neurons with fully recurrent connections.

are evolved for 5000 generations. Each genotype is a vector of 98 real values (85 synaptic connections, 5 decay constants and 8 bias terms) that are initially chosen uniformly random from the range $[-10, 10]$. Subsequent generations are produced by a combination of selection with elitism and mutation. Recombination is not used. At every generation, the best 20 genotypes are selected for reproduction, and each generates 4 offspring. The genotypes of the selected parents are copied in the subsequent generation; the genotypes of the 4 offspring are mutated with a 50% probability of adding a random Gaussian offset $N(0, 1)$ to each real-valued gene. During evolution, genotype parameters are constrained to remain within the range $[-10, 10]$. They are mapped to produce CTRNN parameters with the following ranges: connection weights $W(j, i) \in [-4, 4]$; biases $\beta \in [-4, 4]$; concerning decay constants, the genetically encoded parameters are first mapped onto the range $[-1, 3]$ and then exponentially mapped onto $\tau \in [10^{-1}, 10^3]$. The lower bound of τ corresponds to the integration step size used to update the controller; the upper bound is arbitrarily chosen and it is bigger than the maximum length of a trial.

7.2.1.4 The fitness function

During the evolution, a genotype is mapped into a control structure that is cloned and downloaded onto all the *s-bots* taking part to the experiment (i.e., we use a homogeneous group of *s-bots*). Groups of 3 *s-bots* are evaluated 10 times—i.e., 10 trials, 5 performed in *Env.A* and 5 in *Env.B*. Each trial lasts 65 seconds and differs from the others in the initialisation of the random number generator, which influences mainly the *s-bots* starting positions and orientations and the *way out* amplitude, if present.

The evaluation function takes into account the behavioural state in which the *s-bots* should be and it rewards their movements accordingly. When *s-bots* are placed in *Env.A*, they should search for and traverse the *way out*, therefore they should always be in state $\mathcal{S}$. When *s-bots* are placed in *Env.B*, they should initially search for the *way out*, being in state $\mathcal{S}$, and after some searching they should aggregate, therefore switching to state $\mathcal{C}$. In order to evaluate the behaviour in *Env.B*, we ignore the time needed for searching the *way out* and we consider that an *s-bot* switches to state $\mathcal{C}$ when it encounters the circular band for the first time. In this way, we can systematically evaluate the movements of an *s-bot* according to its behavioural state:

$$\mathcal{S}(s,t) = \mathit{Env.A}\ \mathbf{OR}\ d_M(s,t) < 1.0, \qquad \mathcal{C}(s,t) = \mathbf{NOT}\ \mathcal{S}(s,t), \tag{7.8}$$

where $d_M(s,t)$ is the maximum distance from the centre reached by *s-bot* s at time t. In other words, an *s-bot* is considered to be in state $\mathcal{S}$ if it is placed in *Env.A* or if the maximum distance it reached from the centre of the arena is smaller than one meter, which corresponds to the inner radius of the circular band. Otherwise, an *s-bot* is considered to be in state $\mathcal{C}$. Having defined the behavioural states at time t, an *s-bot* s should maximise its distance from the centre of the arena when in state $\mathcal{S}$, while it should minimise its distance from the centre of mass of the group when in state $\mathcal{C}$. Therefore, for each *s-bot* s at step t, we compute the measure $d(s,t)$ according to the following equation:

$$d(s,t) = \begin{cases} ||\mathbf{X}(s,t) - \mathbf{X}_o|| & \text{if } \mathcal{S}(s,t), \\ 1.0 - ||\mathbf{X}(s,t) - \mathbf{X}_c(t)|| & \text{if } \mathcal{C}(s,t), \end{cases} \tag{7.9}$$

where $\mathbf{X}(s,t)$ are the coordinates of *s-bot* s at time t, $\mathbf{X}_o$ and $\mathbf{X}_c(t)$ are the coordinates of the centre of the arena and of the centre of mass of the *s-bots*. Therefore, an *s-bot* should always maximise $d(s,t)$ in order to reach the optimal position: in state $\mathcal{S}$, an *s-bot* should move away from the centre, and it is considered successful if it reaches an optimal distance $D_O(\mathcal{S}) = 2.4$ m (i.e., $d(s,t) \geq D_O(\mathcal{S})$); in state $\mathcal{C}$, an *s-bot* should aggregate with the other robots by reducing its distance from the centre of mass of the group, and it is considered successful if it stays below an optimal distance $D_O(\mathcal{C}) = 0.25$ m (i.e., $d(s,t) \geq 1.0 - D_O(\mathcal{C})$). We measure a normalised distance $\tilde{d}(s,t)$ according to the behavioural state as follows:

$$\tilde{d}(s,t) = \begin{cases} \Theta\left(\dfrac{d(s,t)}{D_O(\mathcal{S})}\right) & \text{if } \mathcal{S}(s,t), \\ \\ \Theta\left(\dfrac{d(s,t)}{1.0 - D_O(\mathcal{C})}\right) & \text{if } \mathcal{C}(s,t), \end{cases} \tag{7.10}$$

where $\Theta(x)$ simply bounds the value of x in the interval $[0,1]$. In both behavioural states, $\tilde{d}(s,t) = 1$ indicates that *s-bot* s at least reached the optimal distance D_O at time t. We conventionally say that a successful *s-bot* "achieves the optimal distance D_O".

In order to evolve the desired behaviour, we compute two measures that reward the *s-bot*'s movements both for its absolute position and for the stepwise increment of the $d(s,t)$:

$$f_d(s,t) = \tau_d \cdot f_d(s,t-1) + (1-\tau_d) \cdot \tilde{d}(s,t), \tag{7.11}$$

$$f_i(s,t) = \frac{d(s,t) - d(s,t-1)}{2d_M} + 0.5, \tag{7.12}$$

where $\tau_d = 0.975$ is the time constant of a moving average, and d_M is the maximum distance increment that an *s-bot* can cover in a single simulation cycle. The measure $f_d(s,t)$ rewards the *s-bot* for the absolute position reached, and the moving average is justified by the necessity to reward behaviours that keep the optimal distance for a long time (which also justifies the high value we have chosen for the time constant τ_d). Differently, the measure $f_i(s,t)$ rewards the *s-bot* for the stepwise increments toward an optimal position. Notice that, while in state $\mathcal{S}$ robots should continue to move away from the centre of mass even if they achieved the optimal distance $D_O(\mathcal{S})$, in $\mathcal{C}$ *s-bots* cannot decrease further their distance from the centre of mass once the optimal distance $D_O(\mathcal{C})$ is reached. For this reason, we set $f_i(s,t) = 1.0$ when the *s-bot* is in state $\mathcal{C}$ and $\tilde{d}(s,t) = 1$.

Given the above measures computed for all *s-bots* and for all simulation cycles, the fitness in a trial is computed as follows:

$$F = \frac{1}{N}\sum_{s=1}^{N} f_d(s,T) \cdot \frac{1}{NT}\sum_{s=1}^{N}\sum_{t=1}^{T} f_i(s,t), \tag{7.13}$$

where $N = 3$ is the number of *s-bots* and $T = 650$ is the number of simulation cycles of the trial. Note that a trial is terminated whenever an *s-bot* passes over the black border of the circular band—and in this case $F = 0$—or if *s-bots* collide when in state $\mathcal{S}$.

It is worth mentioning that when computing the individual performance, the behavioural state of an *s-bot* cannot be directly observed, because it is not explicitly encoded in the controller or elsewhere. However, knowing the environment type and looking at the movements of the robot, it is possible to estimate in which state an *s-bot* should be at any given time: when an *s-bot* is placed in *Env.A*, it should search for the *way out* and exit through it, therefore it should be in state $\mathcal{S}$. When an *s-bot* is placed in *Env.B*, it should initially search for the *way out*, being in state $\mathcal{S}$, and at some point it should give up and aggregate, therefore switching to state $\mathcal{C}$. Given that it is not possible to exactly recognise when an *s-bot* switches to state $\mathcal{C}$, we compute the individual performance by considering an *s-bot* in state $\mathcal{C}$ as soon as it encounters the circular band for the first time. Basing on such estimation of the behavioural state, it is possible to systematically evaluate the *s-bot*'s performance.

Note that the evaluation function does not explicitly reward cooperation or communication. It rather rewards those agents that perform the correct movements in each behavioural state, without any reference to the mechanism necessary to switch from one state to the other.

7.2.2 Results

We performed 20 replications of the experiment, most of which were successful. For each evolutionary run, we selected a single controller from the last generation. To do so, we evaluated the 20 best individuals—the *elite* of the last generation—for 200 trials in both environments, and we selected the genotype with the highest average performance. As a result, we obtained 20 controllers—hereafter referred to as $C_1, \ldots, C_{20}$—that were further evaluated for 2000 trials, half in *Env.A* and half in *Env.B*. The obtained results are summarised in table 7.5: in both environments, we computed the average performance and its standard deviation (avg ± std), the rates of success %S (all *s-bots* achieve the optimal distance D_O), failure %F (no *s-bot* achieves the optimal distance D_O), partial success/failure %M (not all *s-bots* are successful or fail) and error %E (*s-bots* collide or cross the black edge of the circular band). In each trial, we also computed the *disparity*, which is defined as the percentage of the circular band that each robot covers in average during a trial: a value smaller than 1 indicates that the single *s-bot* does not search the whole circular band for the *way out*, while a value bigger than 1 indicates that the single *s-bot* performs more than one tour (see figure 7.8). The disparity—together with the success rate—is useful to quantitatively assess the quality of the evolved strategies.[2]

Successful controllers produce good search behaviours when *s-bots* are in state $\mathcal{S}$: *s-bots* avoid collisions and move away from the centre of the arena. Once on the circular band, *s-bots* start looping in search of the *way out*, which is eventually found and traversed when *s-bots* are placed in *Env.A*. On the contrary, if *s-bots* are placed in *Env.B*, the absence of the *way out* is recognised by the *s-bots* through the integration over time of their perceptual flow, which includes the signals that the *s-bots* may emit. As a consequence, a

---

[2] In chapter 4, we defined a measure called offset Δ (see equation 4.2, which is similar to the measure of *disparity*; the difference is that for the calculation of the value of offset Δ we demand that a robot signals, while to measure the coverage of the band when we compute the *disparity*, this is not necessary.

Table 7.5: Post-evaluation results. For both *Env.A* and *Env.B* we compute the performance (avg ± std), the success rate (%S), the rate of partial success/failure (%M), the rate of complete failure (%F) and the error rate (%E). Controllers are grouped according to their classes, as indicated in the first column.

		Env.A					*Env.B*				
		avg ± std	%S	%M	%F	%E	avg ± std	%S	%M	%F	%E
U	c_4	0.82 ± 0.14	92.0	6.5	1.0	0.5	0.37 ± 0.11	19.4	18.9	61.7	0.0
	c_6	0.85 ± 0.06	98.6	1.2	0.0	0.2	0.31 ± 0.08	0.9	30.6	68.4	0.1
	c_{14}	0.83 ± 0.15	91.3	6.2	0.0	2.5	0.46 ± 0.15	2.5	65.1	24.0	8.4
	c_{17}	0.66 ± 0.07	74.3	25.4	0.1	0.2	0.39 ± 0.08	4.9	78.8	16.3	0.0
B	c_1	0.86 ± 0.11	97.7	0.8	0.0	1.5	0.69 ± 0.07	95.9	2.8	1.3	0.0
	c_5	0.85 ± 0.13	92.1	5.7	0.0	2.2	0.57 ± 0.14	66.8	16.9	16.1	0.2
	c_8	0.83 ± 0.15	90.3	7.6	0.4	1.7	0.57 ± 0.12	34.3	55.2	9.2	1.3
	c_{10}	0.88 ± 0.07	99.0	0.6	0.0	0.4	0.66 ± 0.07	94.1	2.1	3.7	0.1
	c_{16}	0.85 ± 0.14	94.4	4.1	0.0	1.5	0.74 ± 0.13	94.1	2.3	1.4	2.2
M	c_3	0.83 ± 0.15	85.8	11.7	0.0	2.5	0.63 ± 0.09	87.6	8.1	3.4	0.9
	c_7	0.79 ± 0.20	89.3	5.5	0.0	5.2	0.62 ± 0.25	49.5	34.2	10.5	5.8
	c_{11}	0.86 ± 0.07	98.9	0.6	0.0	0.5	0.61 ± 0.07	87.6	9.5	2.7	0.2
	c_{13}	0.85 ± 0.09	94.3	5.2	0.0	0.5	0.62 ± 0.07	93.0	5.3	0.8	0.9
	c_{19}	0.81 ± 0.15	94.8	2.3	0.6	2.3	0.67 ± 0.12	91.7	3.8	1.9	2.6
	c_{20}	0.87 ± 0.06	99.6	0.0	0.0	0.4	0.59 ± 0.07	79.3	11.3	9.3	0.1
C	c_2	0.86 ± 0.10	98.6	0.1	0.0	1.3	0.82 ± 0.12	97.1	0.4	0.9	1.6
	c_9	0.87 ± 0.08	99.2	0.0	0.0	0.8	0.78 ± 0.12	88.1	8.3	3.1	0.5
	c_{12}	0.87 ± 0.05	99.6	0.3	0.0	0.1	0.74 ± 0.11	87.8	6.4	5.4	0.4
	c_{15}	0.86 ± 0.08	99.3	0.0	0.0	0.7	0.78 ± 0.13	96.6	0.4	0.6	2.4
	c_{18}	0.84 ± 0.18	95.8	0.0	0.0	4.2	0.83 ± 0.17	95.3	0.3	1.0	3.4

behavioural transition from state $\mathcal{S}$ to state $\mathcal{C}$ can be observed. The modalities with which the transition is performed vary significantly across the different solutions synthesised during different evolutionary runs. However, looking at the behaviour produced by the evolved controllers, we recognised some similarities that let us classify the controllers in 4 classes.

Class $\mathbf{U} = \{C_4, C_6, C_{14}, C_{17}\}$ encompasses the "unsuccessful" controllers, that is, those controllers that solve the task only in part. These controllers generally produce appropriate search behaviours when *s-bots* are in state $\mathcal{S}$, as confirmed by the good performance and the high success rate in *Env.A* (see table 7.5). However, when *s-bots* are placed in *Env.B*, they fail in systematically aggregating, scoring a low performance and a poor success rate.

The second class $\mathbf{B} = \{C_1, C_5, C_8, C_{10}, C_{16}\}$ consists of controllers that produce a strategy which we call "bouncing" after the aggregation behaviour of the *s-bots* in state $\mathcal{C}$: *s-bots* search for each other by continuously bouncing off the circular band, so that they sooner or later meet and remain close. Communication is not exploited,[3] and consequently

---

[3] Only C_{16} exploits signalling to trigger a synchronous switch to state $\mathcal{C}$.

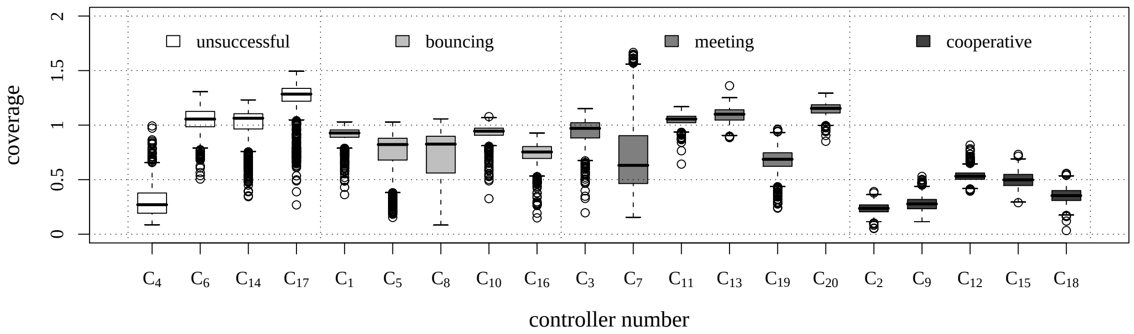


Figure 7.8: The *disparity* of the evolved controllers. Boxes represent the inter-quartile range of the data, while the horizontal lines inside the boxes mark the median values. The whiskers extend to the most extreme data points within 1.5 times the inter-quartile range from the box. The empty circles mark the outliers.

each *s-bot* individually switches from state $\mathcal{S}$ to state $\mathcal{C}$, without any reference to the state of the other robots. The bouncing behaviour is resilient to possible individual failures in *Env.A*: by bouncing off the circular band, *s-bots* can continue searching for the *way out*, even if less efficiently. This corresponds to high success rates in *Env.A* despite that the *s-bots* perform in average less than one tour over the circular band, as indicated by the corresponding disparity (see figure 7.8).

The third class $\mathbf{M} = \{C_3, C_7, C_{11}, C_{13}, C_{19}, C_{20}\}$ encompasses controllers that produce a strategy which we refer to as "meeting", due to the fact that *s-bots* aggregate by encountering at a meeting point, which is normally close to the centre of the arena. Except for C_7 and C_{19}, controllers of this class do not make use of communication. The main difference with class **B** controllers resides in the aggregation behaviour, which lets robots leave the band and move in circles close to the centre of the arena, waiting for the other *s-bots* to reach a similar position. This behaviour is not robust with respect to possible decision errors in *Env.A*. As a consequence, evolution shaped the controllers of this class to be characterised by a high disparity (see figure 7.8): *s-bots* perform more than one loop over the circular band before switching to state $\mathcal{C}$, which corresponds to robust individual decisions and a high success rate in *Env.A*.

The last class $\mathbf{C} = \{C_2, C_9, C_{12}, C_{15}, C_{18}\}$ is named "cooperative" because it encompasses controllers that produce communicative behaviours exploited for cooperation in the decision making. In fact, *s-bots* are able to share the information they collect over time through their signalling behaviour. The *s-bots* initially emit a sound signal, and they stop only after looping on the circular band for some time. If any robot finds the *way out*, signalling continues, inducing all other *s-bots* to remain in state $\mathcal{S}$ and to keep searching for the *way out*. This leads to a high success rate in *Env.A*, and no complete failures are observed (see table 7.5). When the *way out* is not present, all robots eventually stop signalling, allowing the transition to state $\mathcal{C}$ and triggering the aggregation behaviour. By sharing the information through communication, *s-bots* can collectively search the circular band, splitting the task among them: as shown by the disparity data in figure 7.8, each *s-bot* covers from a quarter to half circle when placed in *Env.B*. This allows to consistently reduce the search time, achieving high performance and high success rates. Communication is fundamental here, because it provides robustness to the decision-making process and it makes the system more efficient by reducing the time necessary to take the decisions to switch from solitary to collective behaviours.

In order to quantitatively compare the performance of the behaviours produced by the evolved controllers, we used the performance data recorded over 2000 trials to perform a series of pairwise Wilcoxon tests among all possible controller couples, which allowed to produce the following ranking:

$$\begin{gathered} C_4 \prec C_6 \prec C_{17} \prec C_{14} \prec C_3 \prec C_8 \prec \{C_{13}, C_{11}\} \prec C_{19} \prec C_1 \prec \\ \prec C_{20} \prec C_{10} \prec C_5 \prec C_7 \prec \{C_{16}, C_{12}\} \prec C_{15} \prec C_9 \prec C_2 \prec C_{18}, \end{gathered}$$

where $C_i \prec C_j$ indicates that C_j is statistically better than C_i with 99% confidence. Controllers that have no statistical difference are reported in curly brackets. All class **U** controllers have a low rank, as one would expect. Instead, it is worth noting that class **C** controllers perform statistically better than the others. Moreover, other controllers making use of communication but with a different strategy (namely C_7-Meeting and C_{16}-Bouncing) occupy a good position in the rank. We can conclude that communication can improve the efficiency and the robustness of the decision-making process. Robots exploiting only local interactions are prone to decision errors or to behaviours that are less efficient. Therefore, by cooperating through communication *s-bots* increase their ability to make correct and unanimous decisions, consequently achieving a better performance.

7.2.3 Conclusion

We have studied the decision-making mechanisms that can let a group of robots switch from a solitary to a collective behaviour. We have faced the problem through an evolutionary approach in order to limit the a priori assumptions and search broadly the space of the possible solutions. The results we obtained demonstrate that suitable decision-making mechanisms can be evolved. Moreover, by providing the robots with a simple communication channel, the evolved cooperative strategies display higher efficiency and enhanced robustness of the system. The use of communication generally results in a faster and more robust decision-making process—a result that was also obtained in the experiment detailed in chapter 4. Communication increases the otherwise limited information available to each robot, not only about the quality of the physical environment but also and above all about the social environment and about the internal states of other robots that, by definition, are not directly accessible.

Future work will address a systematic analysis of the evolutionary pressures that shaped the above mechanisms. Also, further testing with real robots has to be performed in order to assess the portability of the evolved controllers on the real *s-bot*. Finally, we plan to combine the decision-making processes studied here with the evolutionary machinery that brought forth self-assembly, as described in detail in chapter 5. Our aim is to produce the first example of functional self-assembly of real *swarm-bots* based on completely evolved controllers. Of course, it will probably be unlikely that we can employ the controllers for assembling robots already developed as a “plug-in”. Most probably, we will have to launch new evolutionary processes to obtain assembly in a new, *functional* context. If the assembly will be functional to the accomplishment of a certain task and will also depend on the environmental structures, we might expect the development of different role allocation mechanisms than those observed in chapter 5.

Chapter 8

General conclusions

In this final section, we draw general conclusions from our experimental work. We start by giving a brief summary of our contributions in section 8.1; this leads us to section 8.2, where we discuss directions for evolutionary robotics research that we consider interesting and promising.

8.1 A brief summary of our contributions

In this thesis we have presented original research work in which an automatic process was employed to design controllers for groups of autonomous robots. More specifically, artificial evolution was used to set the parameters of dynamical neural networks directly controlling the robots' actuators and receiving input from the robots' sensory system.

Our contributions were on the design of non-trivial individual and collective behaviour; in particular, our contributions include:

- The evolution of social behaviour and communication in a task where communication was not a priori required. Communicative and non-communicative behaviour co-evolved from scratch and time-based categorisation skills proved to be precursors of evolved signals.

- The successful porting of time-dependent neuro-controllers on real robots for a collective task requiring the integration over time of the robots' perceptions.

- The evolution of self-assembly between two autonomous robots and the demonstration on real hardware of the effectiveness of the evolved neuro-controllers in directly controlling the actuators of two agents. We prove that coordination and allocation of roles in self-assembling robots can be achieved without the use of explicit communication or complex coordination mechanisms.

- The first example of homogeneous control for morphologically heterogeneous agents, using evolved dynamical neural networks, which specialise "on-the-fly" on the basis of each robot's distinct morphology.

Our approach narrows down the engineer's assumptions about how a task should be solved, about how robots should perceive their world. Solutions found in this way can exploit subtle environmental features as they are perceived through the robot's sensors. Better controllers than those hand-crafted by human developers can be obtained in some cases, since the automatic process can detect features of the robot-environment interaction space that may be adaptive or beneficial for the robots. It is important to point out that these features may not always be a priori evident to the experimenter. Moreover, the approach exploits the interaction among robots and between robots and environment in order to define the quality of a given controller, which may significantly contribute to the adaptivity of the controllers. The adaptivity of a robotic system might be reduced should the decisions concerning solitary or social behaviour be a priori determined by the experimenter.

Our experiments focus on non-reactive collective tasks, that is, tasks where some sort of memory is required on behalf of the robots in order to successfully carry them out and the individual or collective decision-making is temporal. Arguably, the study of such tasks can contribute to the complexification of tasks attacked by evolutionary techniques. Moreover, in chapter 7, section 7.2, we have presented some initial experiments on the evolution of functional self-assembly, a complex task that requires the robots engaged to display complex categorisation skills, fine sensory-motor coordination and cooperation at various different levels. A criticism often addressed to ER is that it fails to efficiently address complex engineering tasks (Matarić and Cliff, 1996). The description of ER in Wikipedia contains the following comment:[1]

> Lately, the difficulty in "scaling up" the complexity of the robot tasks has shifted attention somewhat towards the theoretical end of the field rather than the engineering end.

The author of this thesis agrees that the average complexity achieved by evolutionary techniques is lower than that achieved by other design techniques, behaviour-based robotics for example. However, we argue that the experiments presented in this thesis contribute to scaling up the complexity tackled by ER. Furthermore, the experiments had concrete engineering objectives as well as theoretical objectives, since the tasks we considered were real-world collective tasks.

Also, for two of the tasks considered, namely self-assembly (see chapter 5) and social co-operative categorisation (see chapter 4), we have showed that the evolved neuro-controllers proved to be extremely efficient when controlling a group of real robots. This also contributes to strengthening the evidence that the ER approach is a trustworthy technique for the control of real robots. Again quoting Wikipedia we get:

> Transferring controllers evolved in simulation to physical robots is very difficult and a major challenge in using the ER approach. The reason is that evolution is free to explore all possibilities to obtain a high fitness, including any inaccuracies of the simulation. This need for a large number of evaluations,

---

[1] See `http://en.wikipedia.org/wiki/Evolutionary_robotics`.

> requiring fast yet accurate computer simulations, is one of the limiting factors of the ER approach.

The author of this thesis fully agrees that the issue of the properly designed simulation is extremely important in ER and that the existing examples of successful porting of evolved neuro-controllers to real robots are still rather few. However, we have shown that by modelling only the relevant issues of the robots' environment accurately and by avoiding a globally accurate and very precise simulation, we can achieve perfect results in real-world conditions.

8.2 Looking forward

From an engineering problem-solving perspective, we believe that evolutionary techniques will be used in the future to tackle complex tasks. The future might bring robots with a wide variety of complex sensors whose readings may need to be combined in a very sophisticated way so to produce sensible robot behaviour. For example, a robotic hand of the future might contain hundreds of tactile sensors, and the precise and efficient control of actuators present in it—on the basis of the current sensory readings but also of their history—can be a very delicate and challenging task. Perhaps automatic approaches are the right tool to address such issues.

Shifting to the other, more theoretical end, we believe that ER can be a valuable tool at the disposal of biologists when studying issues as animal or human communication and interaction. The research work presented in this thesis can be seen as a foundation to study complex communication phenomena in natural and artificial organisms. Here, among other research questions, we tried to identify pre-requisites for communication. According to the author, a question that should be tackled in the future with the ER approach is how we can really obtain complex communication skills (or even language skills) in a population of initially non-communicating agents. Succeeding in answering this question would definitely be relevant and useful in the effort to answer one of the harder scientific problems currently: the evolutionary origins of human language.

Moreover, this approach could definitely complement the huge existing literature on game theory, trying to explain one of the major transitions in evolution, the evolution of communication. Using ER, we could revisit the evolution of cooperation and altruistic communication in nature. Consider, for example, the prisoner's dilemma: agents provided with certain cognitive skills have typically signalling capabilities built-in. The ER approach could evolve signalling capabilities from scratch, shedding light on the origin of communicative behaviours. Interestingly, the community is moving into this very promising and exciting direction (see Lipson, 2007), and the first results prove that while deceptive communication arises in a purely competitive setting, cooperative communication arises only subject to group or kin selection (see Floreano et al., 2007).

To conclude, in the eyes of the author, ER is an approach that lies in the delicate balance between theoretical investigation and real-world engineering applications. We believe that it is profitable for both sides if knowledge circulates and is shared: for example, the theoretical challenge of identifying the minimal conditions for self-assembly could be very beneficial in an engineering setup. Also, theoretical assumptions could be verified or revisited upon incoming real-world evidence. Perhaps the ER approach can be combined with other approaches, depending on the goals of the experimenter. For engineering purposes, hybrid approaches (partly hand-coded, partly evolved) that adhere to the principles of autonomous decision-making and that aim at the adaptiveness of the system might prove to work well and to be efficient. For more theoretical questions, blending the ER philosophy with, for example, game theory, evolutionary dynamics or complex systems theory could release huge hidden potential and allow us to study complex phenomena at the core of (artificial) life from a more fine-grained point of view.

Bibliography

Ampatzis, C., Tuci, E., Trianni, V., and Dorigo, M. (2005). Evolving communicating agents that integrate information over time: A real robot experiment. In Talbi, E.-G., Liardet, P., Collet, P., Fonlupt, C., and Lutton, E., editors, *7th International Conference on Artificial Evolution, EA'05*, pages 1–12, Lille, France. University of Lille.

Ampatzis, C., Tuci, E., Trianni, V., and Dorigo, M. (2008). Evolution of signaling in a multi-robot system: Categorization and communication. *Adaptive Behavior*, 16(1):5–26.

Anderson, C., Theraulaz, G., and Deneubourg., J.-L. (2002). Self-assemblages in insect societies. *Insectes Sociaux*, 49(2):99–110.

Anthes, N., Putz, A., and Michiels, N. (2006). Gender conflicts, sex role preferences and sperm trading in hermaphrodites: a new framework. *Animal Behaviour*, 72:1–12.

Arkin, R. C. (1998). *Behavior-based robotics*. MIT Press, Cambridge, MA.

Balch, T. and Arkin, R. C. (1994). Communication in reactive multiagent robotic systems. *Autonomous Robots*, 1(1):27–52.

Balch, T. and Arkin, R. C. (1998). Behavior-based formation control for multiagent robot teams. *IEEE Transactions on Robotics and Automation*, 14(6):926–939.

Baldassarre, G., Nolfi, S., and Parisi, D. (2003). Evolving mobile robots able to display collective behaviour. *Artificial Life*, 9:255–267.

Baldassarre, G., Trianni, V., Bonani, M., Mondada, F., Dorigo, M., and Nolfi, S. (2007). Self-organised coordinated motion in groups of physically connected robots. *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics*, 37(1):224–239.

Baronchelli, A., Felici, M., Loreto, V., Caglioti, E., and Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics*. P06014.

Beer, R. (1995). On the dynamics of small continuous-time recurrent neural networks. *Adaptive Behavior*, 3(4):471–511.

Beer, R. (2003a). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11(4):209–243.

Beer, R. (2006). Parameter space structure of continuous-time recurrent neural networks. *Neural Computation*, 18:3009–3051.

Beer, R. and Gallagher, J. (1992). Evolving dynamical neural networks for adaptive behavior. *Adaptive Behavior*, 1(1):91–122.

Beer, R. D. (2003b). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11(4):209–243.

Blynel, J. and Floreano, D. (2003). Exploring the t-maze: Evolving learning-like robot behaviors using CTRNNs. In *Applications of Evolutionary Computing, EvoWorkshops2003: EvoBIO, EvoCOP, EvoIASP, EvoMUSART, EvoROB, EvoSTIM*, volume 2611 of *LNCS*, pages 598–609. Springer Verlag, Berlin, Germany.

Boden, M. (2008). Autonomy: What is it? *Biosystems*, 91:305–308. Special issue on Modelling Autonomy, invited editorial.

Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999). *Swarm Intelligence: From Natural to Artificial Systems.* Oxford University Press, New York, NY.

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

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47:139–159.

Brooks, R. A. (1992). Artificial life and real robots. In Varela, F. J. and Bourgine, P., editors, *Towards a practice of autonomous systems: Proceedings of the* 1^{st} *European Conference on Artificial Life*, pages 3–10. MIT Press, Cambridge, MA.

Brown, H., Weghe, J. V., Bererton, C., and Khosla, P. (2002). Millibot trains for enhanced mobility. *IEEE/ASME Transactions on Mechatronics*, 7:452–461.

Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Theraulaz, G., and Bonabeau, E. (2001). *Self-Organization in Biological Systems.* Princeton University Press, Princeton, NJ.

Campo, A., Nouyan, S., Birattari, M., Groß, R., and Dorigo, M. (2006). Enhancing cooperative transport using negotiation of goal direction. In *Proceedings of the* 18^{th} *Belgium–Netherlands Conference on Artificial Intelligence (BNAIC 2006)*, pages 365–366. University of Namur, Namur, Belgium.

Cao, Y. U., Fukunaga, A. S., and Kahng, A. (1997). Cooperative mobile robotics: Antecedents and directions. *Autonomous Robots*, 4(1):7–27.

Castano, A., Shen, W., and Will, P. (2000). CONRO: Towards deployable robots with inter-robot metamorphic capabilities. *Autonomous Robots*, 8(3):309–324.

Choset, H., Burgard, W., Hutchinson, S., Kantor, G., Kavraki, L. E., Lynch, K., and Thrun, S. (2005). *Principles of Robot Motion: Theory, Algorithms, and Implementation.* MIT Press, Cambridge, MA.

Christensen, A. (2005). Efficient neuro-evolution of hole-avoidance and phototaxis for a swarm-bot. DEA thesis TR/IRIDIA/2005-14, Université Libre de Bruxelles, Bruxelles, Belgium.

Clutton-Brock, T., Albon, S. D., Gibson, R. M., and Guiness, F. E. (1979). The logical stag: adaptive aspects of fighting in red deer (*cervus elaphus l.*). *Animal Behaviour*, 27:211–225.

Couzin, I. (2007). Collective minds. *Nature*, 445:715.

Couzin, I., Krause, J., Franks, N., and Levin, S. (2005). Effective leadership and decision making in animal groups on the move. *Nature*, 433:513–516.

Dall'Asta, L., Baronchelli, A., Barrat, A., and Loreto, V. (2006). Non-equilibrium dynamics of language games on complex networks. *Physical Review E*, 74:036105.

Damoto, R., Kawakami, A., and Hirose, S. (2001). Study of super-mechano colony: concept and basic experimental set-up. *Advanced Robotics*, 15(4):391–408.

Dawkins, R. (1976). *The selfish gene.* Oxford University Press, Oxford, UK.

Dawkins, R. (1986). *The blind watchmaker: Why the evidence of evolution reveals a universe without design.* W. W. Norton, NY.

De Pinedo, M. and Noble, J. (2003). Mind/body problems? Turn to Beer (commentary). *Adaptive Behavior*, 11(4):289–290.

Di Paolo, E. (2000). Behavioral coordination, structural congruence and entrainment in a simulation of acoustically coupled agents. *Adaptive Behavior*, 8(1):27–48.

Di Paolo, E. and Iizuka, H. (2008). How (not) to model autonomous behaviour. *BioSystems*, 91:409–423. Special issue on Modelling Autonomy.

Dorigo, M. (2005). Swarm-bot: A novel type of self-assembling robot. In Murase, K., Sekiyama, K., Kubota, N., Naniwa, T., and Sitte, J., editors, *Proceedings of the 3rd International Symposium on Autonomous Minirobots for Research and Edutainment (AMiRE 2005)*, pages 3–4. Springer-Verlag, Berlin, Germany.

Dorigo, M. and Colombetti, M. (1998). *Robot Shaping: An Experiment in Behavior Engineering.* MIT Press, Cambridge, MA.

Dorigo, M. and Şahin, E. (2004). Guest editorial. Special issue: Swarm robotics. *Autonomous Robots*, 17(2–3):111–113.

Dorigo, M. and Stützle, T. (2004). *Ant Colony Optimization.* MIT Press, Cambridge, MA.

Dorigo, M., Trianni, V., Şahin, E., Groß, R., Labella, T. H., Baldassarre, G., Nolfi, S., Deneubourg, J.-L., Mondada, F., Floreano, D., and Gambardella, L. M. (2004). Evolving self-organizing behaviors for a swarm-bot. *Autonomous Robots*, 17(2–3):223–245.

Dudek, G. and Jenkin, M. (2000). *Computational principles of mobile robotics.* Cambridge University Press, Cambridge, UK.

Elgar, M. A. (1987). Food intake rate and resource availability: flocking decisions in house sparrows. *Animal Behaviour*, 35:1168–1176.

Ellis, C. (1966). Agonistic behaviour in the male starling. *The Wilson Bulletin*, 78(2):208–224.

Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14:179–211.

Fernández, F., Borrajo, D., and Parker, L. E. (2005). A reinforcement learning algorithm in cooperative multi-robot domains. *Journal of Intelligent and Robotic Systems*, 43(2-4):161–174.

Fernandez-Juricic, E., Smith, R., and Kacelnik, A. (2005). Increasing the costs of conspecific scanning in socially foraging starlings affects vigilance and foraging behaviour. *Animal Behaviour*, 69:73–81.

Floreano, D., Mitri, S., Magnenat, S., and Keller, L. (2007). Evolutionary conditions for the emergence of communication in robots. *Current Biology*, 17:514–519.

Fong, T., Nourbakhsh, I., and Dautenhahn, K. (2002). A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3-4):143–166.

Fukuda, T. and Nakagawa, S. (1987). A dynamically reconfigurable robotic system (concept of a system and optimal configurations). In *Proceedings of the 1987 IEEE International Conference on Industrial Electronics, Control and Instrumentation*, pages 588–595. IEEE Computer Society Press, Los Alamitos, CA.

Fukuda, T., Nakagawa, S., Kawauchi, Y., and Buss, M. (1988). Self organizing robots based on cell structures - CEBOT. In *Proceedings of the 1988 IEEE Int. Workshop on Intelligent Robots*, pages 145–150. IEEE Computer Society Press, Los Alamitos, CA.

Fukuda, T. and Ueyama, T. (1994). *Cellular Robotics and Micro Robotic Systems.* World Scientific Publishing, London, UK.

Funahashi, K. and Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, 6(6):801–806.

Garnier, S., Gautrais, J., and Theraulaz, G. (2007). The biological principles of swarm intelligence. *Swarm Intelligence*, 1(1):3–31.

Gibson, J. J. (1977). The theory of affordances. In Shaw, R. and Bransford, J., editors, *Perceiving, Acting and Knowing. Toward an Ecological Psychology*, chapter 3, pages 67–82. Lawrence Erlbaum Associates, Hillsdale, NJ.

Gigliotta, O. and Nolfi, S. (2008). On the coupling between agent internal and agent/ environmental dynamics: Development of spatial representations in evolving autonomous robots. *Adaptive Behavior*, 16(2–3):148–165.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning.* Addison-Wesley, Reading, MA.

Groß, R., Bonani, M., Mondada, F., and Dorigo, M. (2006a). Autonomous self-assembly in swarm-bots. *IEEE Transactions on Robotics*, 22(6):1115–1130.

Groß, R. and Dorigo, M. (2008a). Evolution of solitary and group transport behaviors for autonomous robots capable of self-assembling. *Adaptive Behavior.* In press.

Groß, R. and Dorigo, M. (2008b). Self-assembly at the macroscopic scale. *Proceedings of the IEEE.* Accepted for publication.

Groß, R., Dorigo, M., and Yamakita, M. (2006b). Self-assembly of mobile robots—from swarm-bot to super-mechano colony. In *Proceedings of the* 9^{th} *International Conference on Intelligent Autonomous Systems*, pages 487–496. IOS Press, Amsterdam, The Netherlands.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7:1–52.

Harvey, I., Di Paolo, E., Tuci, E., and Wood, R. (2005). Evolutionary Robotics: A New Scientific Tool for Studying Cognition. *Artificial Life*, 11(1-2):79–98.

Hasson, O. (1994). Cheating signals. *Journal of Theoretical Biology*, 167:223–238.

Hauser, M. D. (1997). *The evolution of communication.* MIT Press/BradfordBooks.

Hirose, S. (2001). Super mechano-system: New perspectives for versatile robotic systems. In Rus, D. and Singh, S., editors, *Proceedings of the* 7^{th} *International Symposium on Experimental Robotics, (ISER)*, volume 271 of *Lecture Notes in Control and Information Sciences*, pages 249–258. Springer-Verlag, Berlin, Germany.

Hirose, S., Shirasu, T., and Fukushima, E. (1996). Proposal for cooperative robot "Gunryu" composed of autonomous segments. *Robotics and Autonomous Systems*, 17:107–118.

Hockett, C. (1960). The origin of speech. *Scientific American*, 203:88–96.

Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge, UK.

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

Hölldobler, B. and Wilson, E. O. (1978). The multiple recruitment systems of the african weaver ant, oecophylla longinoda (latreille) (hymenoptera: Formicidae). *Behavioural Ecology and Sociobiology*, 3:19–60.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558.

Izzo, D. and Pettazzi, L. (2007). Autonomous and distributed motion planning for satellite swarm. *Journal of Guidance Control and Dynamics*, 30(2):449–459.

Jakobi, N. (1997). Evolutionary robotics and the radical envelope of noise hypothesis. *Adaptive Behavior*, 6:325–368.

Kaelbling, L. P., Littman, M. L., and Moore, A. W. (1996). Reinforcement learning: a survey. *Journal of Artificial Intelligence Research*, 4:237–285.

Kandel, E., Schwartz, J., and Jessell, T. (2000). *Principles of neural science.* McGraw-Hill/Appleton and Lange, New York, NY, 4^{th} edition.

Keinan, A., Sandbank, B., Hilgetag, C., Meilijson, I., and Ruppin, E. (2006). Axiomatic scalable neurocontroller analysis via the shapley value. *Artificial Life*, 12:333–352.

Kelly, I. and Keating, D. (1996). Flocking by the fusion of sonar and active infrared sensors on physical autonomous mobile robots. In *Proceedings of the* 3^{rd} *International Conferences on Mechatronics and Machine Vision In Practice*, pages 1–4. IEEE Computer Society Press, Los Alamitos, CA.

Kube, C. R. and Bonabeau, E. (2000). Cooperative transport by ants and robots. *Robotics and Autonomous Systems*, 30:85–101.

Kube, C. R. and Zhang, H. (1997). Task modelling in collective robotics. *Autonomous Robots*, 4:53–72.

Labella, T. H., Dorigo, M., and Deneubourg, J.-L. (2006). Division of labour in a group of robots inspired by ants' foraging behavior. *ACM Transactions on Autonomous and Adaptive Systems*, 1(1):4–25.

LaValle, S. M. (2006). *Planning algorithms.* Cambridge University Press, Cambridge, UK.

Lewis, D. (1969). *Convention: A philosophical study.* Harvard University Press, Cambridge, MA.

Lioni, A., Theraulaz, G., and Deneubourg, J.-L. (2001). The dynamics of chain formation in œcophylla longinoda. *Journal of Insect Behavior*, 14:679–696.

Lipson, H. (2007). Evolutionary robotics: Emergence of communication. *Current Biology*, 17:R330–R332.

Liu, Z., Ang Jr., M. H., and Seah, W. K. G. (2005). Reinforcement learning of cooperative behaviors for multi-robot tracking of multiple moving targets. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, pages 1289–1294. IEEE Press, Piscataway, NJ.

Livnat, A. and Pippenger, N. (2006). An optimal brain can be composed of conflicting agents. *PNAS*, 103(9):3198–3202.

Maas, W. and Bishop, M. B. (1999). *Pulsed neural networks*. MIT Press, Cambridge, MA.

MacLennan, B. and Burghardt, G. (1993). Synthetic ethology and the evolution of cooperative communication. *Adaptive Behavior*, 2(2):161–188.

Marocco, D., Cangelosi, A., and Nolfi, S. (2003). Evolutionary robotics experiments on the evolution of language. *Philosophical Transactions of the Royal Society of London - A*, 361(1811):2397–2421.

Marocco, D. and Nolfi, S. (2006a). Emergence of communication in embodied agents evolved for the ability to solve a collective navigation problem. *Connection Science*, 19(1):53–74.

Marocco, D. and Nolfi, S. (2006b). Self-organization of communication in evolving robots. In Rocha, L., Yaeger, L., Bedau, M., Floreano, D., Goldstone, R., and Vespignani, A., editors, *Proceedings of the* 10^{th} *International Conference on the Simulation and Synthesis of Living Systems (Artificial Life X)*, pages 178–184. MIT Press, Cambridge, MA.

Matarić, M. J. (1997a). Learning social behavior. *Robotics and Autonomous Systems*, 20:191–204.

Matarić, M. J. (1997b). Reinforcement learning in the multi-robot domain. *Autonomous Robots*, 4(1):73–83.

Matarić, M. J. and Cliff, D. (1996). Challenges in evolving controllers for physical robots. *Robotics and Autonomous Systems*, 19(1):67–83.

Maynard-Smith, J. (1982). *Evolution and the theory of games*. Cambridge University Press, Cambridge, UK.

Maynard-Smith, J. and Harper, D. (2003). *Animal signals*. Oxford University Press, Oxford, UK.

Maynard-Smith, J. and Price, G. (1973). The logic of animal conflict. *Nature*, 246(5427):15–18.

Miglino, O., Lund, H., and Nolfi, S. (1995). Evolving mobile robots in simulated and real environments. *Artificial Life*, 2(4):417–434.

Millor, J., Halloy, J., Deneubourg, J.-L., and Amé, J.-M. (2006). From individual discrimination capabilities to collective decision making. *Journal of Theoretical Biology*, 239:313–323.

Mirolli, M. and Parisi, D. (2008). How producer biases can favor the evolution of communication: An analysis of evolutionary dynamics. *Adaptive Behavior*, 16(1):27–52.

Mitchell, T. (1997). *Machine learning.* McGraw-Hill Education (ISE Editions), Columbus, OH, 1^{st} edition.

Mondada, F. and Floreano, D. (1995). Evolution of neural control structures: some experiments on mobile robots. *Robotics and Autonomous Systems*, 16(2-4):183–195.

Mondada, F., Gambardella, L. M., Floreano, D., Nolfi, S., Deneubourg, J.-L., and Dorigo, M. (2005). The cooperation of swarm-bots: Physical interactions in collective robotics. *IEEE Robotics & Automation Magazine*, 12(2):21–28.

Mondada, F., Pettinaro, G. C., Guignard, A., Kwee, I. V., Floreano, D., Deneubourg, J.-L., Nolfi, S., Gambardella, L. M., and Dorigo, M. (2004). Swarm-bot: A new distributed robotic concept. *Autonomous Robots*, 17(2–3):193–221.

Noble, J. (1998). Tough guys don't dance: intention movements and the evolution of signalling in animal contests. In *Proceedings of the 5^{th} International conference on simulation of adaptive behavior (From animals to animats 5)*, pages 471–476. MIT Press, Cambridge, MA.

Noble, J. (1999). Cooperation, conflict and the evolution of communication. *Adaptive Behaviour*, 7(3–4):349–370.

Noble, J. (2000a). Cooperation, competition and the evolution of prelinguistic communication. In Knight, C., Studdert-Kennedy, M., and Hurford, J., editors, *The Emergence of Language*, pages 40–61. Cambridge University Press, Cambridge, UK.

Noble, J. (2000b). Talk is cheap: Evolved strategies for communication and action in asymmetrical animal contests. In *Proceedings of the 6^{th} International conference on simulation of adaptive behavior (From animals to animats 6)*, pages 481–490. MIT Press, Cambridge, MA.

Nolfi, S. (2002). 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(2):231–244.

Nolfi, S. (2005). Emergence of communication in embodied agents: co-adapting communicative and non-communicative behaviours. *Connection Science*, 17(3-4):231–248.

Nolfi, S. and Floreano, D. (2000). *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines.* MIT Press, Cambridge, MA.

Nolfi, S. and Marocco, D. (2001). Evolving robots able to integrate sensory-motor information over time. *Theory in Biosciences*, 120:287–310.

Nouyan, S., Campo, A., and Dorigo, M. (2008a). Path formation in a robot swarm. self-organized strategies to find your way home. *Swarm Intelligence*, 2(1):1–23.

Nouyan, S., Groß, R., Bonani, M., Mondada, F., and Dorigo, M. (2006). Group transport along a robot chain in a self-organised robot colony. In *Proceedings of the* 9*th* *International Conference on Intelligent Autonomous Systems*, pages 433–442. IOS Press, Amsterdam, The Netherlands.

Nouyan, S., Groß, R., Bonani, M., Mondada, F., and Dorigo, M. (2008b). Teamwork in self-organised robot colonies. *IEEE Transactions on Evolutionary Computation*. Accepted for publication.

Nowak, M. (2006). Five rules for the evolution of cooperation. *Science*, 314:1560–1563.

Nowak, M. and Sigmund, K. (1998). The dynamics of indirect reciprocity. *Journal of Theoretical Biology*, 194:561–574.

O'Grady, R., Groß, R., Mondada, F., Bonani, M., and Dorigo, M. (2005). Self-assembly on demand in a group of physical autonomous mobile robots navigating rough terrain. In Capcarrere, M., Freitas, A., Bentley, P., Johnson, C., and Timmis, J., editors, *Proceedings of the* 8*th* *European Conference on Artifcial Life (ECAL05)*, number 3630 in LNCS, pages 272–281. Springer-Verlag, Berlin, Germany.

Paine, R. W. and Tani, J. (2005). How hierarchical control self-organizes in artificial adaptive systems. *Adaptive Behavior*, 13(3):211–225.

Parker, E. (1998). ALLIANCE: An Architecture for Fault Tolerant Multirobot Cooperation. *IEEE Transactions on Robotics and Automation*, 14:220–240.

Parker, L. E. (2002). Distributed algorithms for multi-robot observation of multiple moving targets. *Autonomous Robots*, 12(3):231–255.

Quinn, M. (2001). Evolving communication without dedicated communication channels. In *Advances in Artificial Life:* 6*th* *European Conference on Artificial Life (ECAL01)*, number 2159 in LNCS, pages 357–366. Springer-Verlag, Berlin, Germany.

Quinn, M., Smith, L., Mayley, G., and Husbands, P. (2002). Evolving teamwork and role allocation with real robots. In *Artificial Life VII: Proceedings of the* 8*th* *International Conference on the Simulation of Adaptive Behavior (SAB'02)*, pages 302–311. MIT Press, Cambridge, MA.

Quinn, M., Smith, L., Mayley, G., and Husbands, P. (2003). Evolving controllers for a homogeneous system of physical robots: Structured cooperation with minimal sensors. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, 361:2321–2344.

Rubenstein, M., Payne, K., Will, P., and Shen, W.-M. (2004). Docking among indepenpent and autonomous CONRO self-reconfigurable robots. In *Proceedings of 2004 IEEE International Conference on Robotics and Automation (ICRA'04)*, volume 3, pages 2877–2882. IEEE Computer Society Press, Los Alamitos, CA.

Santos, F. C., Pacheco, J. M., and Lenaerts, T. (2006). Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proceedings of the National Academy of Science*, 103(9):3490–3494.

Seyfarth, R., Cheney, D., and Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, 28(4):1070–1094.

Sherman, P. W. (1977). Nepotism and the evolution of alarm calls. *Science*, 197:1246–1253.

Spears, W., Spears, D., Hamann, J., and Heil, R. (2004). Distributed, physics-based control of swarms of vehicles. *Autonomous Robots*, 17(2–3):137–162.

Steels, L. (1999). *The Talking Heads experiment. Volume 1. Words and meanings.* VUB Artificial Intelligence Laboratory (1999), Antwerpen, Belgium.

Steels, L. (2007). The recruitment theory of language origins. In Lyon, C., Nehaniv, C. L., and Cangelosi, A., editors, *Emergence of Language and Communication*, pages 129–151. Springer-Verlag, Berlin, Germany.

Støy, K. (2001). Using situated communication in distributed autonomous mobile robots. In *Proceedings of the 7th Scandinavian conference on Artificial Intelligence*, pages 44–52. IOS Press, Amsterdam, The Netherlands.

Struhsaker, T. T. (1967). Auditory communication among vervet monkeys (cercopithicus aethiops). In Altmann, S. A., editor, *Social Communication Among Primates.* University of Chicago Press, Chicago, IL.

Sutton, R. S. and Barto, A. G. (1988). *Introduction to reinforcement learning.* MIT Press, Cambridge, MA.

Szamado, S. and Szathmary, E. (2006). Selective scenarios for the origin of natural language. *Trends in Ecology and Evolution*, 21:555–561.

Tang, F. and Parker, L. (2005). Asymtre: Automated synthesis of multi-robot task solutions through software reconfiguration. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA'05)*, pages 1501–1508. IEEE Computer Society Press, Los Alamitos, CA.

Tinbergen, N. (1952). The curious behavior of the stickleback. *Scientific American*, 187:22–26.

Tinbergen, N. (1953). *The herring gull's world.* Collins, London, UK.

Tinbergen, N. (1964). The evolution of signalling devices. In Etkin, W., editor, *Social Behavior and Organization Among Vertebrates*, pages 206–230. University of Chicago Press, Chicago, IL.

Trianni, V. and Dorigo, M. (2005). Emergent collective decisions in a swarm of robots. In Arabshahi, P. and Martinoli, A., editors, *Proceedings of the 2005 IEEE International Symposium on Swarm Intelligence (SIS 2005)*, pages 241–248. IEEE Press, Piscataway, NJ.

Trianni, V. and Dorigo, M. (2006). Self-organisation and communication in groups of simulated and physical robots. *Biological Cybernetics*, 95:213–231.

Trianni, V., Nolfi, S., and Dorigo, M. (2006). Cooperative hole avoidance in a swarm-bot. *Robotics and Autonomous Systems*, 54(2):97–103.

Trianni, V., Tuci, E., and Dorigo, M. (2004). Evolving functional self-assembling in a swarm of autonomous robots. In Schaal, S., Ijspeert, A., Billard, A., Vijayakamur, S., Hallam, J., and Meyer, J.-A., editors, *From Animals to Animats 8. Proceedings of the* 8^{th} *International Conference on Simulation of Adaptive Behavior (SAB 04)*, pages 405–414. MIT Press, Cambridge, MA.

Tuci, E., Groß, R., Trianni, V., Mondada, F., Bonani, M., and Dorigo, M. (2006). Cooperation through self-assembly in multi-robot systems. *ACM Transactions on Autonomous and Adaptive Systems*, 1(2):115–150.

Tuci, E., Quinn, M., and Harvey, I. (2002). An evolutionary ecological approach to the study of learning behaviour using robot-based model. *Adaptive Behavior*, 10(3-4):201–221.

Tuci, E., Trianni, V., and Dorigo, M. (2004). 'Feeling' the flow of time through sensory-motor coordination. *Connection Science*, 16(4):301–324.

Turgut, A., Celikkanat, H., Gokce, F., and Şahin, E. (2008). Self-organised flocking in mobile robot swarms. *Swarm Intelligence*, 2(2–3). In press.

Urzelai, J. and Floreano, D. (2001). Evolution of adaptive synapses: Robots with fast adaptive behavior in new environments. *Evolutionary Computation*, 9:495–524.

Vicentini, F. and Tuci, E. (2006). *Swarmod*: a 2d *s-bot*'s simulator. Technical Report TR/IRIDIA/2006-005, IRIDIA, CoDE, Université Libre de Bruxelles. This paper is available at `http://iridia.ulb.ac.be/IridiaTrSeries.html`.

Vicentini, F. and Tuci, E. (2007). Scalability in evolved neurocontrollers that guide a swarm of robots in a navigation task. In *Proceedings of the* 2^{nd} *International Workshop on Swarm Robotics*, volume 4433 of *LNCS*, pages 206–220. Springer-Verlag, Berlin, Germany.

V.Jones, C. and Matarić, M. J. (2006). Behavior-based coordination in multi-robot systems. In Ge, S. and Lewis, F., editors, *Autonomous Mobile Robots: Sensing, Control, Decision-Making, and Applications,*. CRC Press, Boca Raton, FL.

Von Frisch, K. (1967). *The dance language and orientation of bees.* Harvard University Press, Cambridge, MA.

Webb, B. (2000). What does robotics offer animal behaviour? *Animal Behaviour*, 60:545–558.

Werner, G. and Dyer, M. (1992). Evolution of communication in artificial organisms. In Langton, C., Taylor, C., Farmer, D., and Rasmussen, S., editors, *Artificial Life II*, pages 659–687, Redwood City, CA. Addison-Wesley Pub.

Whitesides, G. and Grzybowski, B. (2002). Self-assembly at all scales. *Science*, 295:2418 – 2421.

Williams, G. C. (1966). *Adaptation and natural selection.* Princeton University Press, Princeton, NJ.

Wischmann, S. and Pasemann, F. (2006). The emergence of communication by evolving dynamical systems. In *Proceedings of the* 9^{th} *International Conference on Simulation of Adaptive Behaviour (From Animals to Animats 9)*, volume 4095 of *LNAI*, pages 777–788. Springer-Verlag, Berlin, Germany.

Yamakita, M., Taniguchi, Y., and Shukuya, Y. (2003). Analysis of formation control of cooperative transportation of mother ship by smc. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA'03)*, volume 1, pages 951–956. IEEE Computer Society Press, Los Alamitos, CA.

Yamauchi, B. M. and Beer, R. D. (1994). Integrating reactive, sequential and learning behavior using dynamical neural networks. In *Proceedings of the* 3^{rd} *international conference on simulation of adaptive behavior (From animals to animats 3)*, pages 382–391. MIT press, Cambridge, MA.

Yim, M., Zhang, Y., and Duff, D. (2002a). Modular robots. *IEEE Spectrum*, 39(2):30–34.

Yim, M., Zhang, Y., Roufas, K., Duff, D., and Eldershaw, C. (2002b). Connecting and disconnecting for chain self-reconfiguration with PolyBot. *IEEE/ASME Transactions on Mechatronics*, 7(4):442–451.

Ziemke, T. and 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.