Self-organised path formation in a swarm of robots

Valerio Sperati · Vito Trianni · Stefano Nolfi

Abstract In this paper, we study the problem of exploration and navigation in an unknown environment from an evolutionary swarm robotics perspective. In other words, we search for an efficient exploration and navigation strategy for a swarm of robots, which exploits cooperation and self-organisation to cope with the limited abilities of the individual robots. The task faced by the robots consists in the exploration of an unknown environment in order to find a path between two distant target areas. The collective strategy is synthesised through evolutionary robotics techniques, and is based on the emergence of a dynamic structure formed by the robots moving back and forth between the two target areas. Due to this structure, each robot is able to maintain the right heading and to efficiently navigate between the two areas. The evolved behaviour proved to be effective in finding the shortest path, adaptable to new environmental conditions, scalable to larger groups and larger environment size, and robust to individual failures.

Keywords Evolutionary robotics · Swarm robotics · Self-organisation · Path formation

V. Sperati · V. Trianni · S. Nolfi
Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche, via San Martino della Battaglia 44, 00185 Rome, Italy
url: http://laral.istc.cnr.it

V. Sperati
e-mail: valerio.sperati@istc.cnr.it

V. Trianni
e-mail: vito.trianni@istc.cnr.it

S. Nolfi
e-mail: stefano.nolfi@istc.cnr.it

V. Trianni (✉)
IRIDIA-CoDE, ULB, Avenue F. Roosevelt, n. 50, CP 194/6, 1050 Brussels, Belgium
url: http://iridia.ulb.ac.be
e-mail: vtrianni@ulb.ac.be

1 Introduction

Exploration and navigation in unknown environments represent basic activities for most animal species, and efficient strategies can make the difference between death and survival. For this reason, Nature presents a wide range of possibilities, each particularly adapted to the task to be accomplished and to the sensory-motor and cognitive abilities of the species under observation. In primates, as well as in other animals, navigation abilities are usually linked to mental representations of the environment, referred to as 'cognitive maps'. For instance, it has been found that specific neurons of the rodent hippocampus (called 'place cells') have a high firing rate in correspondence of specific locations in the environment (O'Keefe and Nadel 1978). Neural representations seem to also characterise the behaviour of insects. A map-like organisation of spatial memory has been proposed for honeybees, which are able to retrieve the navigation path on the basis of learned landmarks around the hive (Menzel et al. 2005). A similar strategy is also employed by the desert ants of the genus *Cataglyphis*. These ants couple the landmark-based strategy with their skylight polarisation compass and path integrator (i.e., ants integrate over time the path covered through a sort of vector summation). This strategy allows them to return to the nest following a straight line (Wehner 2003). Ant species that forage in groups rely on a collective strategy for exploration and navigation, exploiting the well-known mechanism of pheromone trail formation: when moving from a foraging patch to the nest, ants lay a blend of pheromones that can be exploited by other ants to reach the same patch. Thanks to this strategy, ants can efficiently navigate in the environment and optimise the path between nest and food (Goss et al. 1989; Detrain and Denebourg 2009).

In robotics too, much attention has been paid to the exploration and navigation problems, and several different strategies have been proposed. Probabilistic approaches are employed to solve the simultaneous localisation and mapping problem (SLAM, see Thrun 2003; Bailey and Durrant-Whyte 2006). Map-based representations are also implemented with biologically inspired approaches (Filliat and Meyer 2003; Hafner 2005; Gigliotta and Nolfi 2008). Similarly, landmark-based navigation and path integration have been exploited, often with a close look at biology (Zeil et al. 2009; Lambrinos et al. 1997; Vickerstaff and Di Paolo 2005). When multiple robots are contemporaneously present, the exploration and navigation strategies can leverage on the collective effort. Maps can be built by putting together different pieces of information collected by different robots (Burgard et al. 2005; Thrun and Liu 2005; Pfingsthorn et al. 2008). Also localisation through odometry and path integration can improve thanks to the shared effort of multiple robots exchanging structured information (Rekleitis et al. 2001; Martinelli et al. 2005). In a swarm robotics context, however, the solution of the exploration and navigation problem needs to take into account the limited abilities of the individual robots, which are often characterised by local sensing and communication (Dorigo and Şahin 2004). In this context, the focus is rather on the self-organising process that leads the group as a whole to overcome the individual limitations and to present an overall efficient behaviour.

In this paper, we take an evolutionary perspective to the synthesis of efficient exploration and navigation strategies for a swarm of robots. The goal is to find novel interaction and co-operation modalities that can cope with the limited abilities of the single robots. For this purpose, we design the swarm behaviour through evolutionary robotics techniques (Nolfi and Floreano 2000; Floreano et al. 2008), which proved to be particularly useful to synthesise self-organising collective behaviours characterised by properties such as robustness, flexibility and scalability (Trianni 2008; Trianni and Nolfi 2011). The task to be accomplished requires the robots to explore an unknown environment, find two distant target locations,

and efficiently navigate between them. To allow the robots to cooperate and communicate, we provided them with the possibility to visually signal their relative position and orientation by exploiting just two LEDs—a blue and a red one—positioned in their front and rear side, respectively. The analysis of the obtained results indicates that the robots solve the navigation problem by moving between the two areas following a dynamic path formed by the robots themselves. The robots organise in two lines and keep moving between the two areas in opposite directions. This dynamic path forms as a result of the interactions among the robots, mediated by simple rules that regulate how each individual reacts to the local information provided by environmental and social cues. Once this collective structure is formed, it influences the local interactions among the agents in a way that ensures that the path is self-sustained, and that each robot moves directly toward the next target area. In other words, the formation of the dynamic path allows each individual robot to effectively travel back and forth between the two areas and allows the swarm, as a whole, to preserve the information on the location of the two areas. Interestingly, the evolved behaviour presents features that are similar to trail formation in ants, although it is realised through a different mechanism: robots form a trail between the target locations and robustly maintain it, also optimising the shape towards the shortest path. Such effective coordination and communication mechanisms evolve despite the fact that the evolutionary process does not explicitly request to solve the task in a cooperative manner. Additionally, by analysing how the group behaviour generalises with respect to the swarm and environment size, and to the distance between the targets, we observed that the evolved solution is robust and scalable.

The paper is organised as follows. In Sect. 2, we present a brief overview of the exploration and navigation strategies developed in swarm robotics. In Sect. 3, we describe the task in detail and the experimental setup devised to evolve the robot controllers. In Sect. 4, we discuss the obtained results and analyse the evolved behaviour. We also study the generalisation abilities of the evolved behaviour with respect to larger group size and larger distance between the targets, and we compare the swarm performance with a control experiment in which the same behaviour was evolved for a solitary robot. In Sect. 5, we study the dynamics of the collective behaviour introducing a metric based on information entropy. Finally, discussions and conclusions are reported in Sect. 6.

2 Related work

In the swarm robotics domain, the problem of exploration and navigation has been faced with a close look to biological examples, and particularly to mass recruitment through pheromone trails observed in many ant species. However, implementing a robotic system that replicates the ants' foraging behaviour suffers from the complex problem of finding an alternative to the pheromones, given that chemical substances are difficult to exploit in a robotic setup (for some attempts to disperse and sense chemicals, see Russell et al. 1994; Fujisawa et al. 2008). Various approaches have been tested for the purpose of finding alternatives to pheromones. Instead of chemicals, Garnier et al. (2007) used the light of an overhead projector to highlight the passage of small robots, therefore simulating the deposited pheromones: robots used their ambient light sensors to detect the intensity of the simulated pheromone signal. A recent study investigated the use of fluorescent paint to simulate pheromones, given that the fluorescence activates and fades away with similar dynamics (Mayet et al. 2010). RFID tags have also been proposed as devices to implement virtual pheromones. The devices store the pheromone information, and the robots spread and sense such information while passing by (Mamei and Zambonelli 2007).

The need to use special environmental features like a fluorescent painted floor or RFID tags represents a limiting factor for many application scenarios. A different approach is therefore to rely on communication and message passing among robots, which therefore simulate pheromone attributes on a communication network (Payton et al. 2001; Vaughan et al. 2002; Sadat and Vaughan 2010). In these studies, a virtual trail of pheromones is created and followed by the robots. In other studies, the robot themselves are used to form a path between target locations. In some cases, a bucket brigade method is used for transporting items between two locations. The robots therefore form a chain and transfer objects to one another (Drogoul and Ferber 1993; Østergaard et al. 2001). In other studies, robots just act as markers for the trail. They remain static and signal a path between target locations, while explorer robots exploit this path to efficiently navigate. Werger and Matarić (1996) implemented a robotic chain that is maintained through physical contact between neighbouring robots. Nouyan et al. (2008, 2009) devised a path formation algorithm in which robots in the chain communicate through light signals. Two methods are implemented. In the first method, robots signal one of three different colours, forming a cyclic directional pattern that allows them to determine the direction to follow along the robotic chain to reach the 'nest' or the 'goal' locations. In the second method, robots emit a light pattern that indicates the direction towards the 'nest'. Stirling et al. (2010) present a similar study, in which flying robots are employed for exploration and coverage of indoor environments. The robots form graph-like structures by maintaining wireless links between neighbours. Such structures can be exploited by exploring robots to reach different places in the environment. Ducatelle et al. (2011a) demonstrate path formation in heterogeneous swarms of robots, in which wheeled and flying robots cooperate for exploration and navigation in a complex environment.

Other approaches tackle the exploration and navigation problem by taking inspiration from *trophallaxis*, the direct exchange of food items commonly observed in social insects. When implemented in a robotic swarm, robots do not exchange food items but rather exchange the available information about the distance and direction of the target locations (Schmickl and Crailsheim 2008; Gutiérrez et al. 2010). This allows them to rapidly spread through communication the information about the presence of target areas, or to improve the quality of the available information. In a recent work, Ducatelle et al. (2011b) use a similar framework for navigation in unknown environments. When all robots in the swarm have to navigate between two targets, the resulting collective behaviour is similar to the one presented in this paper, although obtained exploiting a more structured and long-range communication.

The above studies are representative of a common methodology in swarm robotics, which starts from a biological example, distills its relevant features and transposes the identified mechanisms in the robotic system. In this respect, our methodology strongly differs, because the self-organising behaviour of the robotic swarm is synthesised through artificial evolution, without a specific biological inspiration (Nolfi and Floreano 2000; Trianni 2008; Trianni and Nolfi 2011). The use of this methodology allows us to discover solutions that might be otherwise difficult to imagine and/or implement by the experimenter, even by taking inspiration from available knowledge on natural behaviour. Indeed, as we will see in Sect. 4, this approach leads to a qualitatively different solution with respect to those described above. A similar approach has been taken by Hauert et al. (2009a), who developed a path formation behaviour for micro air vehicles (MAVs). In this experiment, MAVs are capable of hovering by flying in circles, and form a communication network that extends in the environment, always keeping the connection with the launching station.

Besides the methodological aspects, the studies presented above differ from our approach in many other ways. Unlike the studies that adopt structured communication (Payton et

al. 2001; Vaughan et al. 2002; Sadat and Vaughan 2010; Schmickl and Crailsheim 2008; Gutiérrez et al. 2010; Ducatelle et al. 2011b), we adopt a minimalist approach, using sub-symbolic signalling and neural computation. This minimalist approach adapts well to the evolutionary methodology, and allows the definition of simple control rules that exploit the fine-grained interactions among robots for the emergence of a collective strategy. Moreover, unlike previous work in which robots are dedicated explicitly to signalling the path between the target locations (Werger and Matarić 1996; Nouyan et al. 2008, 2009; Stirling et al. 2010; Ducatelle et al. 2011a), the dynamic path described in this paper involves all robots, which continuously move between the two areas. This solution potentially enables to better exploit the available resources (e.g., robots may transport objects from one area to another while participating in the dynamic path; see also the discussions in Sect. 6). Additionally, the proposed solution does not need to allocate different roles in the swarm, and to define when and how many robots are needed for a specific role.

3 Experimental setup

In this section, we describe the experimental setup designed to evolve efficient navigation in an unknown environment within a swarm robotics context. The goal of each robot is to move back and forth as quickly as possible between two target areas, located within a rectangular arena surrounded by walls. Since no explicit map of the environment is available, and since the robots' sensory range is limited (i.e., targets can be perceived only from a short distance), robots have to explore the environment in order to find the targets and to preserve, in some way, the information concerning the locations of the areas, without relying on continuous time-consuming exploratory actions. We investigate how this task can be solved relying only on a collective strategy that leverages on the coordination and communication abilities of the robots in the swarm.

In order to develop such self-organising behaviour, the robot controller is synthesised through a simple evolutionary algorithm. We show in Sect. 4 that the evolutionary process produces a very interesting collective strategy: the swarm self-organises forming a dynamical structure composed of two rows of robots moving in opposite directions. This dynamical structure connects the targets' locations and allows the agents to navigate between them, overcoming their individual limitations. In the following, we give a detailed description of the experimental setup, we specify the characteristics of the agent and of its controller, we describe the ecological niche in which the robots evolve, and finally we introduce the evolutionary algorithm used to evolve the swarm behaviour.

3.1 The robot and the controller

The experiments have been performed in simulation, using a customised version of *Evorobot*, an open-source software developed for evolutionary robotics experiments.[1] The simulated agent models the *e-puck* robot, a small wheeled robot with a cylindrical body, having a diameter of 7 cm (see Fig. 1(a) and Mondada et al. 2009). The robot has two independent motors controlling two wheels, which provide a differential drive motion and a maximum speed of $\omega_M = 8.2$ cm/s. Additionally, the robot is equipped with eight infrared proximity sensors placed around the body, which allow it to detect obstacles up to a distance
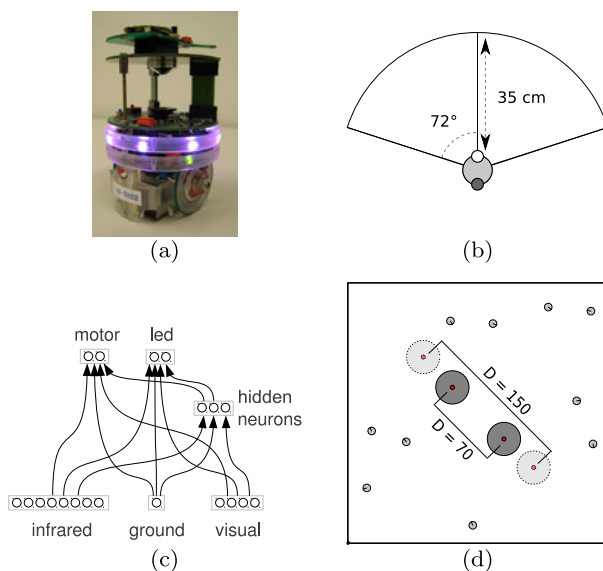
[1]See http://laral.istc.cnr.it/evorobotstar/.

Fig. 1 (**a**) The e-puck (Mondada et al. 2009) is the robot simulated in the experiments. The figure shows the robot equipped with two hardware extensions: the coloured LED communication turret and the omni-directional camera. (**b**) A schematic representation of the robot body and camera sensors. The blue and red LED positions are indicated respectively as a *white* and a *grey dot* on the robot's body. (**c**) The neural network structure, showing the inputs, the hidden layer and the outputs. (**d**) Snapshot of the simulated environment. Ten robots randomly positioned in the environment are represented as *small circles*. The two *grey disks* represent the circular target areas, with a red LED in the centre. The distance between the centres varies from trial to trial in the set $D_{set} = \{70, 90, 110, 130, 150\}$ cm (in the figure, disks are drawn in *dark grey* when $D = 70$ cm, in *light grey* when $D = 150$ cm)

of 2.5 cm. An additional infrared sensor—the ground sensor—is placed beneath the agent, on the front part, allowing it to detect the floor colour: we use it in a binary way to perceive dark areas (the targets), within the white-coloured experimental arena. The robot has a modular architecture that permits to add hardware extensions (*turrets*). In this work, we used the coloured LED turret, which provides signalling abilities through eight RGB LEDs, and the omni-directional vision turret, used to detect the coloured signals emitted by neighbouring robots (see Fig. 1(a) and Floreano et al. 2010). We use only two coloured signals: a blue LED placed in the front of the body, and a red LED placed in the rear (respectively represented as white and grey circles in Fig. 1(b)). The two LEDs can be switched on and off at will, and can be perceived by other robots through the omni-directional camera. We limit the field of view of the robots to two sectors with a width of 72° and a perceptual range of 35 cm, as shown in Fig. 1(b). The robot can detect just the presence of blue or red LEDs within these sectors, for a total of four binary sensory inputs. Notice that the LEDs can be perceived only when the robots are facing them; otherwise they are occluded by the robot's body. In this way, they provide information about the heading of the signalling robot: a blue LED corresponds to the robot's front, a red LED to its rear.

The controller of the robot is a feed-forward neural network, whose structure is displayed in Fig. 1(c). The network is provided with 13 input neurons that relay the normalised sensor values (8 from the proximity sensors, 1 binary ground sensor and 4 binary visual inputs), 3 hidden neurons, and 4 output neurons (2 controlling the angular speed of the wheels, 2 controlling the activation of the red and blue LEDs). The activation O_j of the jth output

neuron is computed as the weighted sum of all input and hidden neurons and a bias term, filtered through a sigmoid function:

$$O_j(t) = \sigma\left(\sum_i w_{ij}^{oI} I_i(t) + \sum_i w_{ij}^{oH} H_i(t) + \beta_j^o\right), \qquad \sigma(z) = \frac{1}{1 + e^{-z}}, \qquad (1)$$

where $I_i(t)$ is the value of the ith input neuron at time t, $H_i(t)$ is the value of the ith hidden neuron at time t, β_j^o is a bias term, and w_{ij}^{oI} and w_{ij}^{oH} are the weights of the synaptic connections, respectively from the input and the hidden neurons. The three internal neurons are leaky integrators, i.e., they maintain a fraction of the previous activation, according to the following equation:

$$H_j(t) = \tau_j H_j(t-1) + (1-\tau_j)\mathcal{H}_j, \qquad \mathcal{H}_j = \sigma\left(\sum_i w_{ij}^{hI} I_i(t) + \beta_j^h\right), \qquad (2)$$

where τ_j is the time constant, w_{ij}^{hI} are the weights of the synaptic connections between input and hidden neurons, and β_j^h are the bias terms of the hidden neurons. All weights and bias terms take values in the range $[-5, 5]$, while the time constants take values in $[0, 1]$. The output neurons are used to control the speed of the wheels, by scaling their value in the range $[-\omega_M, \omega_M]$. LEDs are switched on when the corresponding output neuron crosses the threshold 0.5; otherwise they are switched off. The sensor inputs, the motor outputs and the network internal neurons are updated every 0.1 simulated seconds.

3.2 The environment

The ecological niche in which the robots evolve is a white rectangular arena surrounded by walls (height $H = 250$ cm; width $W \in [250, 290]$ cm; the variable width removes some environmental regularities that could be exploited to solve the exploration problem, and therefore makes the evolution of the behaviour more robust). The arena contains two circular target areas (diameter $d = 32$ cm), each characterised by the dark colour of the floor, and by a red LED placed over the centre (see Fig. 1(d)). This LED is always on, and is indistinguishable from the one provided to the robots. Target areas can be perceived by the robots up to a distance of 35 cm, thanks to the camera sensors. Additionally, a robot can detect being within one of the target areas because of the ground sensor. The target areas are always positioned symmetrically with respect to the arena centre, and the distance D between them is chosen systematically, trial by trial, in the set $D_{set} = \{70, 90, 110, 130, 150\}$ cm. Note that, given these parameters and considering the field of view of the robot cameras, the target areas are never detectable at the same time by a single robot.

3.3 The evolutionary algorithm

The parameters of the neural network controller—connections weights, biases, and time constants—are obtained using artificial evolution (Nolfi and Floreano 2000; Floreano et al. 2008). These parameters are encoded in a binary genotype, using 8 bits for each real number. Evolution works on a population of 100 randomly generated genotypes. After fitness evaluation, the 20 best genotypes survive in the next generation (elitism), and reproduce asexually by generating four copies of their genes with a 3% mutation probability of flipping each bit. The evolutionary process lasts 500 generations.

3.4 Fitness function

In order to evaluate its fitness, a genotype is translated into N identical neural controllers which are downloaded onto N identical robots (i.e., the group is homogeneous, see Floreano et al. (2007) for a discussion on the advantages of genetic relatedness for the evolution of cooperative behaviours). The genotype fitness F is computed by evaluating the behaviour of the robotic group for $M = 15$ trials. All the possible distances between targets are equally experienced, each value in D_{set} being tested for 3 trials. Each trial lasts $T = 6000$ time steps corresponding to 10 simulated minutes. Robots are evaluated only during the second part of a trial ($T_b = 5400$ time steps); during the first part of the trial ($T_a = 600$ time steps) robots can freely move to achieve coordination without contributing to the fitness computation.

The fitness measure we devised rewards the robots for efficiently travelling between the target areas. This measure is based on a simple idea. When a robot arrives in a target area, it virtually loads a fixed amount of energy, which is consumed along the travel proportionally to the robot speed. When a robot arrives at the second target area, the remaining energy is stored, and a new load is assigned to the robot for another travel. To maximise the amount of energy stored, the robot must efficiently navigate between the target areas, therefore choosing the shortest path between the two. Moreover, robots must maximise the number of travels between target areas in the limited time available. The interplay of these two drives—minimising energy consumption and maximising the number of travels—leads to efficient navigation strategies. In fact, long travels are discarded in favour of short ones, in order to save as much energy as possible. However, fast motion on short paths is rewarded in order to maximise the number of travels.

In order to formalise this concept, a robot i is endowed with a virtual energy e_i. In each time step t, the energy level is decreased by a quantity δ_i, dependent on the robot speed:

$$\delta_i(t) = \frac{|\omega_{il}(t)| + |\omega_{ir}(t)|}{2k\omega_M} \tag{3}$$

where ω_{il} and ω_{ir} represent the angular speed of the wheels of robot i, ω_M is the maximum speed, and $k = 400$ is a constant stating that a robot moving at maximum speed consumes one unit of energy in k time steps. The energy level is updated as follows:

$$e_i(t) = \begin{cases} 1 + E_D & \text{if robot } i \text{ enters a new target area} \\ e_i(t-1) - \delta_i(t) & \text{otherwise} \end{cases} \tag{4}$$

Here, $1 + E_D$ is a constant amount of energy provided to the robot when it enters a target area different from the one previously visited, and E_D is the energy that a robot would consume to move in a straight line at maximum speed from one area to the other, that is, what we consider an optimal behaviour. The energy left to the robot when entering a new target area contributes in computing its individual fitness f_i:

$$f_i(t) = f_i(t-1) + \begin{cases} e_i(t-1) & \text{if robot } i \text{ enters a new target area} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

This equation states that for each time step t, if robot i has just entered a target area different from the one previously visited, the remaining energy e_i is added to the fitness f_i.

In this way, a robot displaying optimal behaviour would save exactly a quantity $e = 1$ each time it enters a correct target area, independently of the distance D between the two.

Consequently, f_i would be incremented by 1. At the end of the trial, f_i is scaled with respect to the maximum number of travels that can be performed in T_b time steps:

$$F_i = f_i/f_{\max}, \qquad f_{\max} = r\omega_M \cdot T_b/D_m \qquad (6)$$

where r is the radius of the robot wheel, and $D_m = D - d$ is the minimum distance that must be covered between two target areas. The fitness F of the genotype is computed as the average of the individual fitnesses F_i, computed over all trials:

$$F = \frac{1}{MN} \sum_{m=1}^{M} \sum_{i=1}^{N} F_i. \qquad (7)$$

It is important to notice that this fitness computation does not explicitly reward any coordination or cooperation among the agents to achieve their goal. Nevertheless, as we discuss in the following, the evolved behaviour strongly exploits communication and cooperation.

4 Obtained results and behavioural analysis

In this section, we report the results obtained in ten replications of the experiment, in which we evolved the neural controller for a group of $N = 10$ robots. The obtained results are presented in Sect. 4.1. In Sect. 4.2, we investigate whether the evolved solutions generalise with respect to the distance between the two areas and the number of robots forming the swarm. To verify the extent to which the task can be solved by a single robot and to compare the individual and collective solutions, we performed a control experiment in which the behaviour of a solitary robot is evolved ($N = 1$); the results are reported in Sect. 4.3.

4.1 Collective solution

We performed ten replications of the evolutionary experiment, each starting with a different randomly generated population of genotypes. Each evolutionary run lasted 500 generations. At the end of the evolutionary process, we selected the best genotype of each evolutionary run by choosing among the best solutions of the last 100 generations. To do so, we computed the performance of each of these 100 genotypes by re-evaluating the group behaviour for 500 different trials, and we selected the one with the highest average fitness. A qualitative analysis of the behaviours produced by the best evolved genotypes from the different evolutionary runs revealed that 6 out of 10 result in a good collective exploration and navigation behaviour (roughly corresponding to $F \geq 0.30$). Two runs produced sub-optimal strategies, and two others resulted in unsatisfactory behaviours both at the individual and collective level (see Sperati et al. 2010). The six successful evolutionary runs produced similar collective strategies. In the following, we analyse in detail one of them, namely the one that presents the best generalisation abilities with respect to larger group size and larger distance between the targets,[2] as described in Sect. 4.2. The chosen solution was obtained in evolutionary run 7 and corresponds to the best genotype of generation 488.

---

[2]The genotype that scored the highest average performance belongs to the third evolutionary run and is described in Sperati et al. (2010), but is characterised by worse generalisation abilities with respect to the one chosen in this paper.
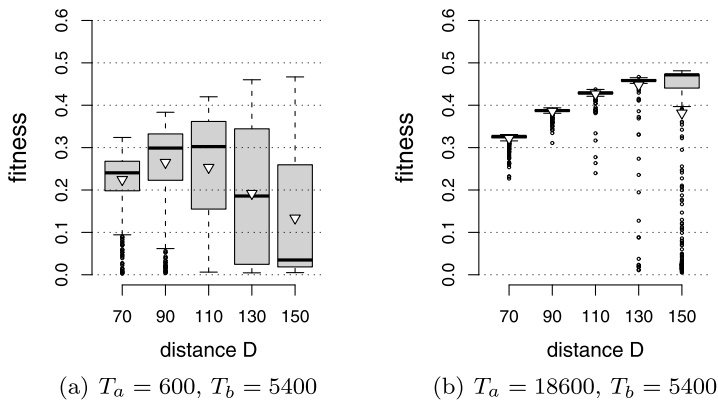
Fig. 2 (**a**) Boxplot of the performances obtained testing the best evolved genotype over 500 standard trials for each distance $D \in D_{set}$ ($T_a = 600$, $T_b = 5400$). *Boxes* represent the inter-quartile range of the data, while the *horizontal lines* inside the boxes mark the median values. The whiskers extend to the most extreme data points within 1.5 times the inter-quartile range from the box. *Circles* mark the outliers. The symbol ▽ indicates the average performance. (**b**) Performance of the best evolved genotype when tested over longer trials ($T_a = 18600$, $T_b = 5400$). A comparison between conditions (**a**) and (**b**) shows how the swarm successfully solves the task (independently of the distance between areas) when enough time is granted

To appreciate the performance of the group, we tested the collective behaviour by systematically varying the distance between the target areas $D \in D_{set}$. The obtained results are presented in Fig. 2(a). We notice that the behaviour seems adapted mostly for an intermediate distance, in which the group scores the highest average performance. With longer distances, the performance across different trials presents a higher variability. This suggests that the group may be able to coordinate in some cases, and not in others. This variability depends on the limited duration of the trial, which is stopped after a fixed number of time steps. In some cases, the trial length is not sufficient for the robots to coordinate, especially in the most difficult cases in which the target areas are farther away. To check this hypothesis, we performed an identical test increasing the duration of the initial coordination period ($T_a = 18600$, $T_b = 5400$ time steps). The results plotted in Fig. 2(b) confirm that for all distances the group attains a good score, which is also very stable across different trials. Moreover, the results indicate that the group behaves better for long distances. In fact, with short distances, the path between the two target areas is overcrowded, and robots interfere with each other, therefore scoring a lower performance. In these conditions, a smaller group behaves better (data not shown).

The visual inspection and the qualitative analysis of the evolved strategy reveal how robots cooperate to efficiently navigate between the target areas. The sequence displayed in Fig. 3 shows how a typical successful trial unfolds in time. Initially robots move independently and explore the environment. In doing so, they signal their relative position and heading to other robots, keeping the front blue LED always switched on. The red LED is switched on when the robots visually detect a blue LED or a red LED on the left visual field. However, in the latter case the red light flashes (i.e., it repeatedly goes on and off). The visual interactions mediated by these signals allow the group to converge to a coherent motion between the two target areas. Eventually, the robots assemble into two rows moving in opposite directions, from one target to the other. We refer to this structured spatio-temporal pattern formed by the robots as a *dynamic chain*. The term *dynamic* well illustrates two interesting features of this structure. First, each robot within the chain is not static, but moves
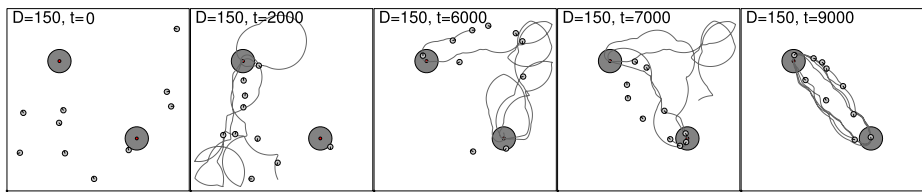
Fig. 3 Temporal sequence recorded in a generic successful trial showing the formation of the *dynamic chain*. The trajectory of one robot for the last 3000 time steps is shown as a *grey line*. See also the video 'robots10-D150.mpeg' ($N = 10$ robots, $D = 150$ cm) in the online supplementary material
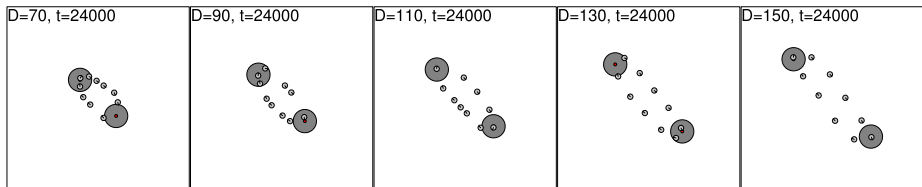

Fig. 4 From *left* to *right*, each snapshot displays the final configuration achieved by $N = 10$ robots at the end of five different trials, in which the target areas are positioned according to the distance values in D_{set}. The shape of the dynamic chain suggests (i) the optimality of the path (a straight line), and (ii) the success of the strategy regardless of the distance D. See also the videos 'robots10-D70.mpeg' ($N = 10$, $D = 70$ cm) and 'robots10-D150.mpeg' ($N = 10$, $D = 150$ cm) in the online supplementary material. For further video footage, refer to http://laral.istc.cnr.it/esm/sperati-etal-2010

continuously along it, swinging between the target areas as requested by the fitness function. Second, the chain connecting the two targets adapts its shape according to the current distance D between areas: the chain direction is optimised, choosing the shortest path between the two areas (a straight line in our setup), and the inter-robot distance varies to fit all robots in the chain, as shown in Fig. 4.

This collective behaviour is the result of simple rules followed by each individual robot and encoded in the neural controller. When a robot has no objects in its perceptual field, it moves clockwise in large circles (with a radius of about 35 cm), the front blue LED always switched on. When a target area is in sight, a robot approaches it in a straight line, making a tight U-turn once reaching it, as a response to the perception of the dark floor. When two robots encounter, they avoid each other by always dodging to the left, exploiting the blue visual signal. This simple action constitutes the basic mechanism for the formation of the dynamic chain: in fact, for a while this interaction straightens the circular path of the agents, and repeated interactions with multiple robots in a row result in an almost straight line motion. As a consequence of mutual influences, a stable path can be created as soon as there is a sufficient number of robots that move in two opposite rows. This can happen when the robots from one row pass to the other. Actually, this is the case when robots perform U-turns at the target areas. In other words, a robot within the chain manages to maintain the right route, not necessarily by exploiting the red signals emitted by the robots in front that move in the same direction, but rather by exploiting the blue signals emitted by the robots belonging to the opposite row. This can be explained by considering that a row of robots moving in the opposite direction provides a cue both of the fact that the next target is located frontally and that a dynamic chain passing between the two target areas has been formed. It is interesting to note that the same mechanism responsible for the formation of the dynamic chain is also responsible for the initial exploration phase. In fact, by avoiding each
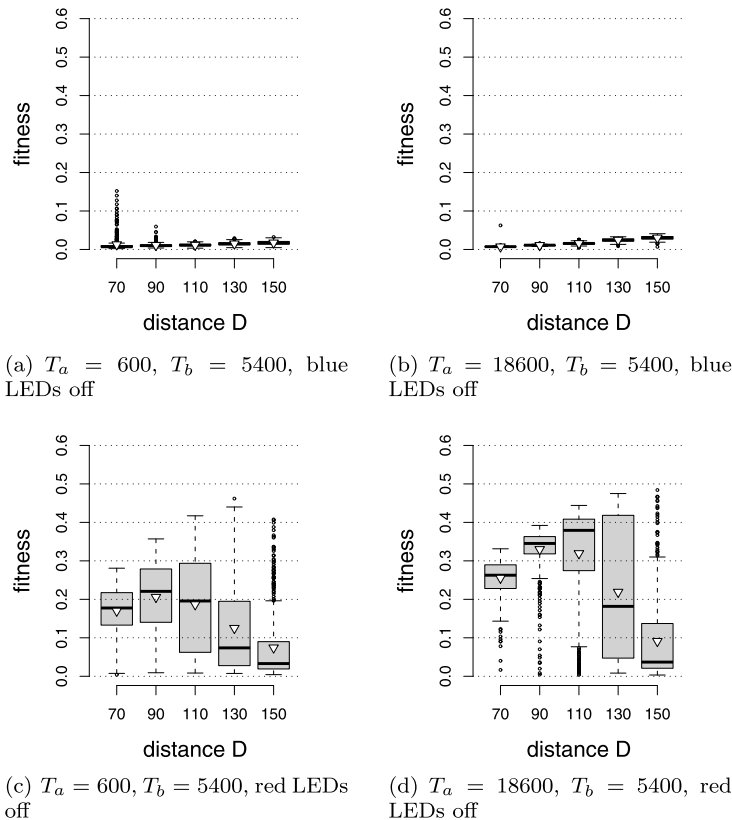
(a) $T_a = 600$, $T_b = 5400$, blue LEDs off

(b) $T_a = 18600$, $T_b = 5400$, blue LEDs off

(c) $T_a = 600$, $T_b = 5400$, red LEDs off

(d) $T_a = 18600$, $T_b = 5400$, red LEDs off

Fig. 5 Performance tests on the role of red and blue signals. Boxes represent the performance over 500 trials. The tests were performed forcing either the blue or the red LEDs off

other exploiting the blue visual signals, robots spread in the arena and explore it thoroughly. From a cognitive point of view, it is possible to speculate that the dynamic chain increases the spatial cognition of the individual robot, as it works as a "collective representation" of the relative direction between the two targets. This information, as already stated, is not directly available given the limited sensory-motor and processing abilities of the individual robot.

As discussed above, the role of the blue LEDs is essential to support the swarm behaviour. To quantitatively assess the role of blue signals, we ran a series of tests in which the blue LEDs were forced off during the whole trial, and measured the corresponding performance. We found that without the use of blue signals, the group is not capable of cooperating to navigate between the two target areas, as indicated by the low performance in Fig. 5(a) and 5(b), respectively for trials with a standard ($T_a = 600$, $T_b = 5400$) and extended duration ($T_a = 18600$). The role of communication through red signals is less definite. Undoubtedly, the red LEDs inside the target areas provide a cue about the direction to follow. However, what is the functional role of the red signals emitted by the robots themselves? To verify the relevance of these signals, we tested the swarm in a control set-up in which the robots' red LEDs are always switched off, during standard and extended duration trials (see Fig. 5(c) and 5(d)). Compared to the data obtained in normal conditions shown in Fig. 2(a) and 2(b), the results obtained suggest that the dynamic chain can still be formed despite the absence
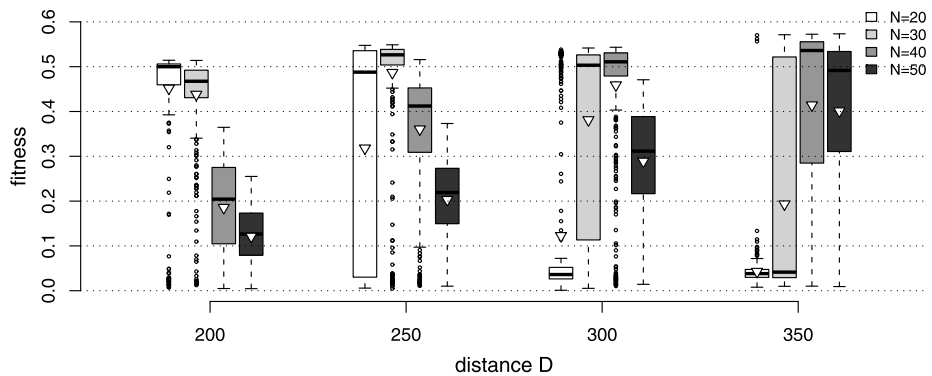
Fig. 6 Generalisation ability for groups of increasing size N and for increasing distance D. Each boxplot corresponds to the performance obtained in 500 trials ($T_a = 36600$, $T_b = 5400$)

of red signals—as indicated by the high performance scored in many trials—but with lower efficiency—as indicated by the higher variability in the scored performance. We suggest that the functional role of the red signals is to accelerate the formation of the dynamic chain and to make the chain more stable. In fact, in standard duration trials the performance is in general lower for all distances, in particular for longer ones (compare Fig. 5(c) with Fig. 2(a)). This means that the formation of the dynamic chain is faster in normal conditions with respect to the system deprived of red signals. Additionally, tests in extended duration trials show that the performance varies considerably with respect to the standard conditions (compare Fig. 5(d) with Fig. 2(b)), in which performance was high and constant. In this case, either chains are not formed at all, or, once formed, they are rather unstable and easily break, especially for longer distances.

4.2 Generalisation abilities

In the previous section, we have described the features of the evolved behaviour, and observed how the system always converges to a dynamic path formation if enough time is granted for coordination. In this section, we test the ability of the system to generalise to different conditions never met during the evolutionary optimisation. In particular, we want to understand whether robots are able to form a path with longer distances and with larger groups. We tested the performance of the group in 16 new conditions, which correspond to 4 group sizes ($N \in \{20, 30, 40, 50\}$ robots) coupled with 4 distances ($D \in \{200, 250, 300, 350\}$ cm). These tests have been performed in a larger arena (fixed height $H = 350$ cm, variable width $W \in [350, 390]$ cm), and in longer trials ($T_b = 5400$, $T_a = 36600$). The quantitative results are shown in Fig. 6.

   We immediately notice that, when the number of robots is sufficiently large, the swarm is successful even when the distance between the two target areas is much wider compared to the conditions experienced during the evolutionary process. With distance $D = 200$ cm, groups with 20 and 30 robots perform best, while larger groups are less efficient. With distance $D = 250$ cm it is possible to notice a similar pattern. However, this time $N = 30$ is the optimal group size. For $D = 300$ cm the size $N = 40$ performs best, while for $D = 350$ cm $N = 40$ and $N = 50$ present good performance, this time with a larger variability. This analysis confirms that the larger the distance between the target areas, the larger the number of robots required to form a stable chain. In fact, as mentioned above, the dynamic chain is
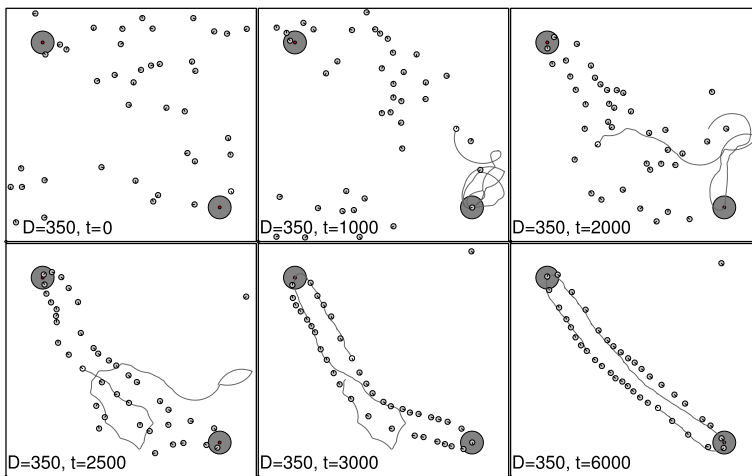
Fig. 7 Six snapshots taken during a successful trial with $N = 40$ robots and $D = 350$ cm. It is possible to notice the chain formation and optimisation through time. The trajectory of a single robot within the group is also shown as a *grey line*. See also the video 'robots40-D350.mpeg' ($N = 40$, $D = 350$ cm) in the online supplementary material. For further video footage on generalisation abilities, refer to http://laral.istc.cnr.it/esm/sperati-etal-2010

maintained as long as there are constantly robots moving in opposite directions uniformly distributed along the path, which implies larger groups for longer distances. The analysis also confirms that a minimum number of robots is necessary to form a path over a certain distance. Similarly, large groups suffer overcrowding when the distance D is too short, as there is no space available to distribute all the robots along the path. However, the dynamic chain can adapt to a wide range of distances. For instance, groups of 30 robots present good performance up to $D = 300$ cm; only with longer distances the performance does systematically drop. The larger number of robots and the longer distances allow us to better appreciate the dynamics of the chain formation, as shown in Fig. 7. It is possible to notice that robots are first attracted around the target areas, where temporary, unstable structures begin to form. Then a stable structure forms connecting the target areas, which afterwards changes shape, optimising the chain direction and the robot positions within the chain.

Large swarms are also robust to individual failures. We have performed a series of tests ($N = 30$, $D = 250$) in which part of the robots are prevented from turning their blue LEDs on, either at the beginning of the trial or after 36600 time steps, when the dynamic chain is presumably already formed. As shown in Fig. 8, the performance starts to deteriorate only when more than 6 out of 30 robots are damaged at the beginning of the trial (white bars in the figure). Even with 10 damaged robots—one third of the group—it is possible to observe a majority of the trials in which the chain is formed, as indicated by the median. In short, the dynamic chain formation is a very robust behaviour with respect to individual failures. This is further confirmed by tests performed by damaging the robots after 36600 time steps (grey bars in Fig. 8). In this condition, even with nearly 50% of damaged robots one can observe that the group is capable of scoring a good performance.

4.3 Individual solution

In order to understand how good the collective strategy is with respect to what can be done individually, we have performed a control experiment in which a neural network controller
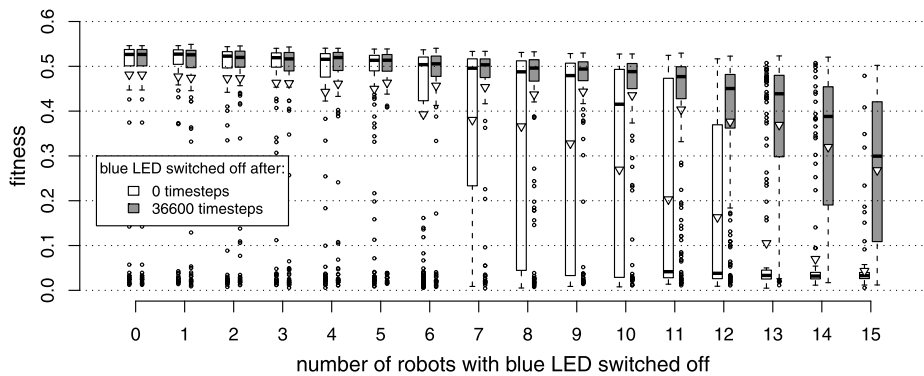
Fig. 8 Group performance when an increasing number of robots are 'damaged' by forcing the blue LED off. Each box represents 200 trials ($T_a = 36600$, $T_b = 5400$, $N = 30$, $D = 250$ cm). The robots are either damaged from the beginning of the trial (*white bars*) or after 36600 time steps, when the chain is presumably already formed (*grey bars*)
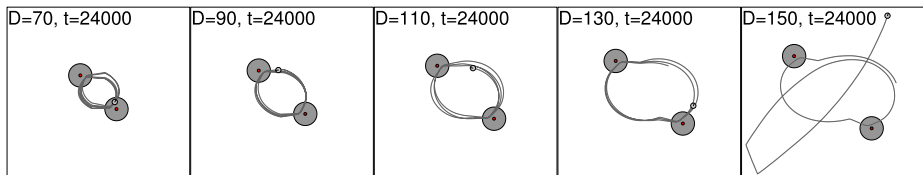

Fig. 9 From *left* to *right*, each snapshot displays the trajectory followed by a single robot in five different trials, in which the target areas are positioned according to the distance values in D_{set}. Notice the failure of the individual strategy when $D = 150$ cm (last frame on the right). See also the video 'single-robot-D150.mpeg' ($N = 1$, $D = 150$ cm) in the online supplementary material. For further video footage about the single robot experiment, refer to http://laral.istc.cnr.it/esm/sperati-etal-2010

is evolved for the same task but with a solitary robot ($N = 1$). We have replicated the experiment in 30 evolutionary runs, each lasting 1000 generations. All other parameters and methodological aspects (i.e., controller architecture, environmental variability, fitness function, evolutionary algorithm) are kept constant. At the end of the evolutionary process, we selected the best evolved genotype among the 30 evolutionary runs for detailed analysis, using the same methodology described in Sect. 4.1. The best evolved genotype in this case corresponds to the one obtained in evolutionary run 27, generation 971.

The evolved behaviour is very simple. At the beginning of a generic trial the robot performs an almost straight trajectory, allowing the robot to explore the arena while avoiding collisions with walls. Once the red LED in a target area is detected, the robot moves towards it, maintaining the beacon between the two visual fields. Then, once the target is reached, the dynamics of the neural network make the robot produce a wide counterclockwise turn. If in doing so it encounters the second target area, the robot performs a further counterclockwise turn that allows it to head back and find, with high probability, the first target area. The trajectories performed in this condition are plotted in Fig. 9. However, this strategy works only when targets are not too distant, and even in these cases the produced trajectories are not straight. Additionally, when $D = 150$, the agent continually gets lost, being unable to maintain the correct route.
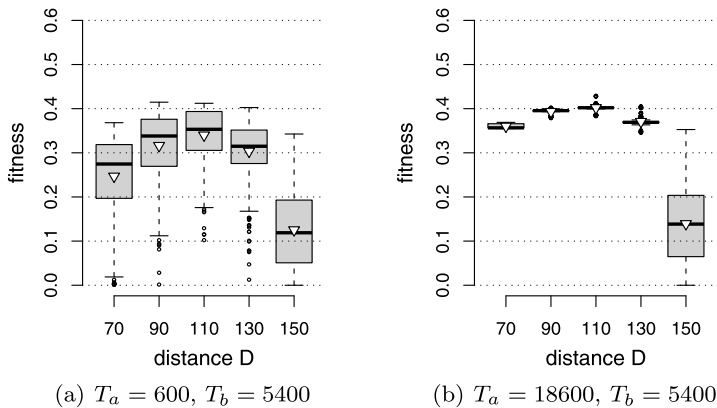
Fig. 10 (**a**) Boxplot of the performances obtained testing the best genotype evolved in the control experiment, in which a single robot is present. Each box represents 500 trials ($T_a = 600$, $T_b = 5400$) for each distance $D \in D_{set}$. (**b**) Performance of the best evolved genotype when tested over longer trials ($T_a = 18600$, $T_b = 5400$). The drop of performance when $D = 150$ indicates that a single agent is not able to solve the task when targets are too far

The quantitative results plotted in Fig. 10(a) confirm the performance drop when target areas are too distant. This performance drop is somehow similar to what we observed with the collective strategy, shown in Fig. 2(a). The differences between the two approaches appear when we test the behaviour in longer trials ($T_a = 18600$, see Fig. 10(b)), in which the extra time granted for the task has evident benefits only for short distances. In this case, it is evident that the individual solution is sub-optimal and properly works only for small distance values. Therefore, we can conclude that the solitary robot cannot produce a navigation strategy as efficient as the one obtained by the group.

5 Dynamics of chain formation

The analyses performed in the previous section showed how the formation of the chain is the outcome of a self-organising process that results solely from the robot-robot interactions. A qualitative analysis of the behaviour suggests that the dynamic chain forms rather abruptly out of a disordered group motion (see, for instance, Fig. 7). In this section, we analyse the process of chain formation, introducing a measure capable of representing with an acceptable approximation the extent to which the chain is formed at every time step. As discussed above, the chain forms and is stable as soon as the group splits into two rows, heading to the two target areas. In this condition, the heading of each robot corresponds approximately to the direction between the two target areas (as shown in the top part of Fig. 11). Therefore, we can use a measure that encodes statistical information about the variability in the heading directions of the robots within the arena, at every time step, to identify whether and when a chain between the two targets is formed.

If we consider the current directions of the N robots as N independent samples of a generic random variable X, we can measure the entropy $H[X]$ (Shannon 1948). This is a statistical measure that describes the probability distribution of this random variable: the entropy is maximised when all the possible values that X can take have the same probability to be observed, while it is null when just one value is systematically observed. According to
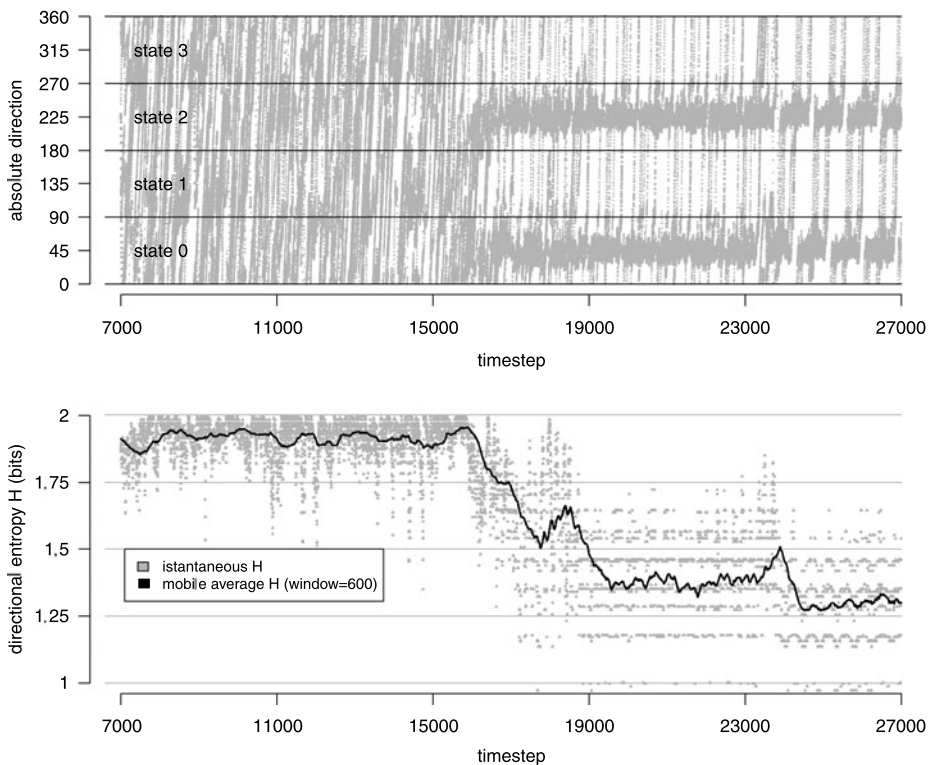
Fig. 11 *Top*: the absolute directions followed by robots is plotted through time (*grey lines*) during a successful trial ($N = 30$, $D = 250$). It is possible to notice how all robots converge and maintain a heading similar to the direction between target areas once the dynamic chain is formed. The figure also shows how the heading direction is discretised in four states for the computation of the entropy $H[X]$. *Bottom*: the respective values of $H[X]$ for each time step are drawn in *grey*, the moving average with a window of 600 time steps is drawn in *black*. The time step when the average drops to values lower than 1.75 corresponds to the formation of the chain

this feature, when the group self-organises in the chain, $H[X]$ has to decrease, because if we sample X in the group, it assumes just two values, corresponding to the two possible heading directions towards the target areas. In order to compute $H[X]$, the robots' heading direction is discretised in $K = 4$ states (see Fig. 11, top), obtaining for each time step N samples of X that are used to estimate the probability distribution. Then, the standard equation for entropy computation is applied:

$$H[X] = -\sum_{k \in K} p(x_k) \cdot \log_2 p(x_k), \tag{8}$$

where X is the observed random variable, k is the current observed state, and $p(x_k)$ is the probability to observe X in the state k. Because of the size of K, we can expect an ideal $H \approx 2$ when the team is not yet organised—i.e., all robots are moving in random directions—and an ideal $H \approx 1$ when the chain is formed—i.e., all robots are moving in just two directions.

The data about one successful trial ($N = 30$, $D = 250$ cm) confirm these expectations. The bottom part of Fig. 11 shows a clear decrease of the entropy H in correspondence of the

polarisation of the robots into two rows (between time step 15000 and 19000), that is, when the chain is likely formed. The value of $H[X]$ is obviously very noisy, because of the very small set of observations per time step ($N = 30$). However, a mobile average with a time window of 600 time steps still returns an appropriate entropy estimation, which highlights the sudden decrease in entropy as soon as the chain is formed. When the entropy decreases and remains below a value of $H = 1.75$, we can reasonably state that the dynamic chain is formed.

Once we empirically established the quality of the delivered measure, we performed an analysis for varying distance D among the target areas, and for the corresponding best group found during the generalisation tests (previously shown in Fig. 6). We recorded 200 successful trials ($T_a = 36600$), selecting those in which the group scored a fitness $F \geq 0.45$, and for each trial we computed the entropy trend. Figure 12(a) shows the results for $N = 30$ and $D = 250$. As we can see, the chain always forms rather abruptly, but in different moments of the trial, due to the randomness of the spatial distribution of the robots.[3] The thick line in the graph represents the average across 200 successful trials in which the chain was formed. This measure is somehow linked to the probability of observing the group in a chain formation: the longer we wait, the higher is the probability that the chain is formed. This is confirmed by comparing the average entropy obtained for different distances D and different group sizes, shown in Fig. 12(b). It is possible to observe that the entropy decay is similar in all cases. We can conclude that the chain formation behaviour is scale invariant, as soon as a sufficient number of robots is available to cover the distance D between the target areas. In fact, we have not observed particular differences in the dynamics of chain formation with varying D and N. Chain formation is not faster for shorter distances than for longer ones (if the proper size of the swarm is chosen), and it happens with the same probability at each time step, no matter what the size of the group is, as clearly shown in Fig. 12(b). This is a very interesting feature for a self-organising artificial system, which further proves its robustness and scalability.

6 Discussion and conclusions

Path formation is a powerful means to collectively produce efficient exploration and navigation strategies. Similarly to pheromone trails in ants, dynamic chains allow a group of robots to efficiently navigate between two target areas. The parallel between pheromone trails and dynamic chains may be overstated, but it is worth noting that the two spatio-temporal structures share common features. For instance, in both cases the stability of the structure is a result of a sustained flux of individuals that support it, in one case by pheromone laying, in the other by coloured signals. Another common feature is the ability to identify the shortest path between two locations. Even though we only performed tests in an obstacle-free arena, we observed that dynamic chains may initially form in curved paths (especially with large groups), which slowly straighten until the shortest route is taken. In future work, we plan to analyse in more detail the evolved behaviour, in order to better understand the properties of the dynamic chain and its relationship with similar behaviours observed in Nature.

It is worth mentioning that the experimental setup we have devised was conceived without explicitly favouring cooperative behaviours. The robots are rewarded on the basis of the efficiency with which each single individual is able to accomplish the task (see Sect. 3.4),

---

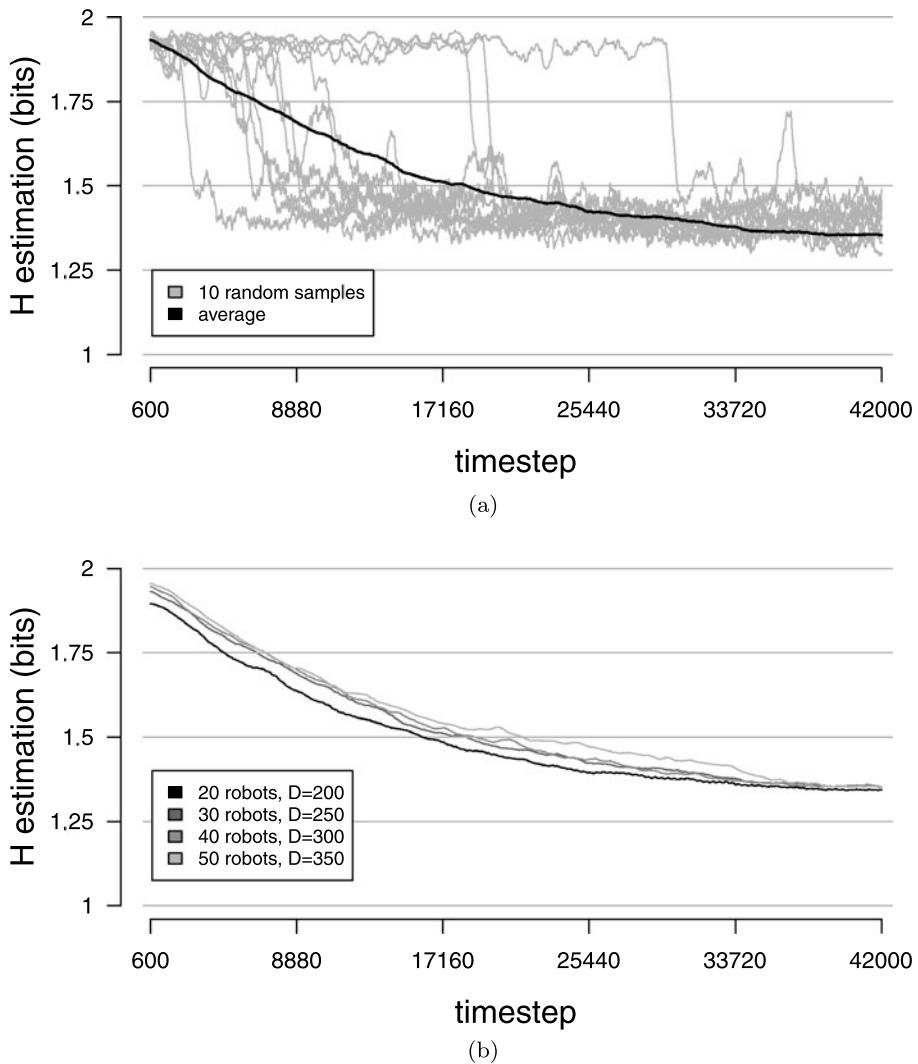[3]The data relative to the other groups are very similar and are not shown.

Fig. 12 (**a**) In *grey*, ten examples of the entropy decay, corresponding to ten successful trials in which $F \geq 0.45$). The *black thick line* corresponds to the average computed over 200 successful trials. (**b**) Average entropy decay for varying distance D and varying group size N

and the evolutionary process may potentially lead to the development of non-cooperative solutions in which being part of a swarm does not provide any advantage. However, implicit selective pressures are present, related to the ecological conditions in which the behaviour evolves. Since the distance between the target areas is larger than the individual perceptual range, robots need to explore the environment to find the target areas. Moreover, in order to navigate from one area to the other without relying on time-consuming exploratory behaviours, the robots should be able to preserve, in some way, the information about the relative location of previously visited areas. Another implicit pressure towards the emergence of coordinated behaviours might arise by the need to reduce the interferences caused

by the other moving individuals. As a consequence, the evolutionary process leads to strategies in which robots coordinate and cooperate to find the target. More specifically, the robots coordinate to form a dynamic chain composed of robots that keep moving back and forth between the two areas along the discovered path. The ability to form dynamic chains allows the robots to preserve, at the level of the group, information concerning the relative positions of previously visited areas, which allows them to achieve a better performance than what can be achieved by individuals operating in isolation. Indeed, the ability to form and maintain a dynamic chain between the two areas allows each individual robot to access local information (constituted by blue and red signals) that provides a straightforward indication on what it has to do in order to navigate directly towards the next area and in order to preserve the chain.

In this process, communication plays a crucial role. In Sect. 4.1, we have described how the formation of the dynamic chain exploits the visual signals exchanged by the robots. The most important role, in this respect, is played by blue LEDs that are always switched on and that are used by the robots to dodge each other on the left. The interaction between the robots mediated by this signal, in fact, tends to lead to the formation of two rows of robots moving in opposite directions. This in combination with other individual skills, such as the ability to turn sharply about 180 degrees once a target area is reached, is one of the key mechanisms that enable the formation and the maintenance of dynamic chains between target areas. One interesting question in this respect concerns how such a communication form emerged evolutionarily (i.e., how the ability to always keep the blue LED on and to react to it by dodging on the left was evolved). A possible answer to this question, that we plan to verify in future studies, is that the tendency to react to blue light by turning to the left is initially evolved in order to reduce the interference between the robots, and is then reused to also produce and maintain dynamic chains (for a related work showing how communication forms can emerge as a side effect of the development of individual behavioural skills, see (De Greef and Nolfi 2010)).

The situation is different for the red signals. We have explicitly designed the experiment to have the robot's red signal mimic the target area. This implies that, as soon as the robot develops an ability to move towards the target areas (i.e., towards the red LED located over the target areas), they will also tend to move towards the red LEDs located on the rear of the other robots. This behavioural bias can facilitate the emergence of a communication skill that consists in the ability of the robots to turn the red LED on when they detect a red LED (according to the receiver bias hypothesis in the evolution of communication by Maynard-Smith and Harper (2003)). The emergence of this form of communication is indeed observed in our experiments. However, the red signal is also turned on when a robot detects a blue LED. The analysis reported in Sect. 4.1 indicates that red signals emitted by the robots speed up the chain formation process and increase the stability of the dynamic chain. However, the limited impact of the presence/absence of this signal on the robots' performance indicates that it plays a minor role. This is probably due to several factors: (i) the interaction mediated by the blue signal is sufficient to lead to the formation of dynamic chains, (ii) the information provided by the red signal is noisy since the direction of the red LED of the robot located ahead might vary, within limits, with respect to the direction of the target area, and (iii) small variations of the orientations of the robot located in line towards a target area tend to reduce the chances that the red signal propagates along the chain through several robots.

It is worth noting that the evolved solution to the swarm exploration and navigation problem cannot be directly applied to different application scenarios, as it was not designed to generalise to different experimental setups (e.g., the presence of large obstacles, the need to perform some work at the target areas such as loading/unloading objects, or the

need to recruit the optimal number of robots in order to build the chain). Nevertheless, the evolved solution could be taken as a reference for engineering swarm robotic systems that present similar dynamics. The evolutionary process actually identified a collective strategy that would have been otherwise difficult to design. The evolved strategy overcomes the limitations of comparable studies: the dynamic chain does not require the simulation of pheromones for the path formation process (Payton et al. 2001; Vaughan et al. 2002; Sadat and Vaughan 2010), and exploits all the available resources without using robots as markers for the chain (Werger and Matarić 1996; Nouyan et al. 2008, 2009; Ducatelle et al. 2011a). Therefore, the dynamics of the discovered process could be reverse-engineered and adapted to different application scenarios (similarly to what was done by Hauert et al. (2009b) with the evolved relay network of MAVs). In this sense, evolutionary robotics can be considered a very useful design method to be combined with more conventional approaches, as it can point to solutions that would be otherwise difficult to obtain (Baldassarre and Nolfi 2009).

In future work, we plan to test the evolved behaviour in hardware exploiting the *foot-bot* robotic platform (Bonani et al. 2010) developed within the European project *Swarmanoid* (grant IST-022888, see http://www.swarmanoid.org). We also aim to test coordinated behaviours among groups of heterogeneous robots. In particular, we plan to exploit the *eye-bot* robotic platform (Roberts et al. 2008) to work as 'smart' target areas. The eye-bot is an aerial robot with the ability to attach to the ceiling, from which it can monitor the environment and detect relevant areas. Foot-bots can perceive an eye-bot only when they are approximately underneath it, due to a camera pointing upward. As a consequence, eye-bots could be exploited by the foot-bots as target areas, and dynamic chains can be formed between them, creating for instance a delivery line from a target to a goal location, which is known by the eye-bots because of their privileged viewpoint. Moreover, eye-bots can move in the environment and exploit the robustness of the dynamic chain to purposely modify its length or its shape. Finally, we will investigate how to form more complex dynamic chains that can trace the shortest path between more than two eye-bots or target areas.

References

Bailey, T., & Durrant-Whyte, H. (2006). Simultaneous localization and mapping: part II. *IEEE Robotics & Automation Magazine*, *13*(3), 108–117.

Baldassarre, G., & Nolfi, S. (2009). Strengths and synergies of evolved and designed controllers: a study within collective robotics. *Journal of Artificial Intelligence*, *173*, 857–875.

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

Burgard, W., Moors, M., Stachniss, C., & Schneider, F.E. (2005). Coordinated multi-robot exploration. *IEEE Transactions on Robotics*, *21*(3), 376–386.

De Greef, J., & Nolfi, S. (2010). Evolution of implicit and explicit communication in a group of mobile robots. In S. Nolfi & M. Mirolli (Eds.), *Evolution of communication and language in embodied agents* (pp. 179–214). Berlin: Springer.

Detrain, C., & Deneubourg, J.-L. (2009). Collective decision and foraging patterns in ants and honeybees. *Advances in Insect Physiology*, *35*, 123–173.

Dorigo, M., & Şahin, E. (2004). Swarm robotics—special issue editorial. *Autonomous Robots*, *17*(2–3), 111–113.

Drogoul, A., & Ferber, J. (1993). From Tom Thumb to the Dockers: Some experiments with foraging robots. In J.-A. Meyer, H. Roitblat, & S. W. Wilson (Eds.), *From animals to animats 2. Proceedings of the second international conference on simulation of adaptive behavior (SAB 92)* (pp. 451–459). Cambridge: MIT Press.

Ducatelle, F., Di Caro, G. A., Pinciroli, C., & Gambardella, L. M. (2011a). Self-organized cooperation between robotic swarms. *Swarm Intelligence*, *5*(2) (this issue).

Ducatelle, F., Di Caro, G. A., Pinciroli, C., Mondada, F., & Gambardella, L. M. (2011b). *Communication assisted navigation in robotic swarms: self-organization and cooperation*. Technical report IDISA-04-11, IDISA, Lugano, Switzerland. Submitted to IROS 2011.

Filliat, D., & Meyer, J.-A. (2003). Map-based navigation in mobile robots—I. A review of localization strategies. *Journal of Cognitive Systems Research*, *4*, 243–282.

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

Floreano, D., Husband, P., & Nolfi, S. (2008). Evolutionary robotics. In B. Siciliano & O. Khatib (Eds.), *Handbook of robotics* (pp. 1423–1451). Berlin: Springer.

Floreano, D., Mitri, S., & Hubert, J. (2010). A robotic platform for studying the evolution of communication. In S. Nolfi & M. Mirolli (Eds.), *Evolution of communication and language in embodied agents* (pp. 303–306). Berlin: Springer.

Fujisawa, R., Imamura, H., Hashimoto, T., & Matsuno, F. (2008). Communication using pheromone field for multiple robots. In *Proceedings of the 2008 IEEE/RSJ international conference on intelligent robots and systems (IROS 2008)* (pp. 1391–1396). New York: IEEE Press.

Garnier, S., Tâche, F., Combe, M., Grimal, A., & Theraulaz, G. (2007). Alice in pheromone land: an experimental setup for the study of ant-like robots. In *Proceedings of the 2007 IEEE swarm intelligence symposium (SIS 2007)* (pp. 37–44). New York: IEEE Press.

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

Goss, S., Aron, S., Deneubourg, J.-L., & Pasteels, J. M. (1989). Self-organized shortcuts in the Argentine ant. *Naturwissenchaften*, *76*, 579–581.

Gutiérrez, A., Campo, A., Monasterio-Huelin, F., Magdalena, L., & Dorigo, M. (2010). Collective decision-making based on social odometry. *Neural Computing & Applications*, *19*(6), 807–823.

Hafner, V. V. (2005). Cognitive maps in rats and robots. *Adaptive Behavior*, *13*, 87–96.

Hauert, S., Zufferey, J.-C., & Floreano, D. (2009a). Evolved swarming without positioning information: an application in aerial communication relay. *Autonomous Robots*, *26*(1), 21–32.

Hauert, S., Zufferey, J.-C., & Floreano, D. (2009b). Reverse-engineering of artificially evolved controllers for swarms of robots. In *Proceedings of the 2009 IEEE congress on evolutionary computation (CEC'09)* (pp. 55–61). New York: IEEE Press.

Lambrinos, D., Kobayashi, H., Pfeifer, R., Maris, M., Labhart, T., & Wehner, R. (1997). An autonomous agent navigating with a polarized light compass. *Adaptive Behavior*, *6*(1), 131–161.

Mamei, M., & Zambonelli, F. (2007). Pervasive pheromone-based interaction with RFID tags. *ACM Transactions on Autonomous and Adaptive Systems*, *2*(2), 1–28.

Martinelli, A., Pont, F., & Siegwart, R. (2005). Multi-robot localization using relative observations. In *Proceedings of the 2005 IEEE international conference on robotics and automation (ICRA 2005)* (pp. 2797–2802). New York: IEEE Press.

Mayet, R., Roberz, J., Schmickl, T., & Crailsheim, K. (2010). Antbots: a feasible visual emulation of pheromone trails for swarm robots. In M. Dorigo, M. Birattari, G. A. Di Caro, R. Doursat, A. P. Engelbrecht, D. Floreano, L. M. Gambardella, R. Groß, E. Şahin, T. Stützle, & H. Sayama (Eds.), *Lecture notes in computer science: Vol. 6234. Proceedings of the 7th international conference on swarm intelligence (ANTS 2010)* (pp. 84–94). Berlin: Springer.

Maynard-Smith, J., & Harper, D. G. (2003). *Animal signals*. London: Oxford University Press.

Menzel, R., Greggers, U., Smith, A., Berger, S., Brandt, R., Brunke, S., Bundrock, G., Hülse, S., Plümpe, T., Schaupp, F., Schüttler, E., Stach, S., Stindt, J., Stollhoff, N., & Watzl, S. (2005). Honey bees navigate according to a map-like spatial memory. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(8), 3040–3045.

Mondada, F., Bonani, M., Raemy, X., Pugh, J., Cianci, C., Klaptocz, A., Magnenat, S., Zufferey, J.-C., Floreano, D., & Martinoli, A. (2009). The e-puck, a robot designed for education in engineering. In P. J. S. Gonçalves, P. J. D. Torres, & C. M. O. Alves (Eds.), *Proceedings of the 9th conference on autonomous*

robot systems and competitions (Vol. 1, pp. 59–65). IPCB: Instituto Politécnico de Castelo Branco, Portugal.

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

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

Nouyan, S., Groß, R., Bonani, M., Mondada, F., & Dorigo, M. (2009). Teamwork in self-organized robot colonies. *IEEE Transactions on Evolutionary Computation*, 13(4), 695–711.

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. London: Oxford University Press.

Østergaard, E. H., Sukhatme, G. S., & Matarić, M. J. (2001). Emergent bucket brigading: a simple mechanism for improving performance in multi-robot constrained-space foraging tasks. In *Proceedings of the fifth international conference on autonomous agents* (pp. 2219–2223). New York: ACM Press.

Payton, D., Daily, M., Estkowski, R., Howard, M., & Lee, C. (2001). Pheromone robotics. *Autonomous Robots*, 11(3), 319–324.

Pfingsthorn, M., Slamet, B., & Visser, A. (2008). A scalable hybrid multi-robot SLAM method for highly detailed maps. In U. Visser, F. Ribeiro, T. Ohashi, & F. Dellaert (Eds.), *Lecture notes in computer science: Vol. 5001. RoboCup 2007: robot soccer world cup XI* (pp. 457–464). Berlin: Springer.

Rekleitis, I., Dudek, G., & Milios, E. (2001). Multi-robot collaboration for robust exploration. *Annals of Mathematics and Artificial Intelligence*, 31, 7–40.

Roberts, J. F., Zufferey, J.-C., & Floreano, D. (2008). Energy management for indoor hovering robots. In *Proceedings of the IEEE/RSJ international conference on intelligent robots and systems (IROS 2008)* (pp. 1242–1247). New York: IEEE Press.

Russell, A., Thiel, D., Deveza, R., & Mackay-Sim, A. (1994). Sensing odour trails for mobile robot navigation. In *Proceedings of the 1994 IEEE international conference on robotics and automation (ICRA'94)* (pp. 2672–2677). New York: IEEE Press.

Sadat, S. A., & Vaughan, R. T. (2010). SO-LOST an ant-trail algorithm for multi-robot navigation with active interference reduction. In H. Fellermann, M. Dorr, M. Hanczyc, L. Ladegaard Laursen, S. Maurer, D. Merkle, P.-A. Monnard, K. Støy, & S. Rasmussen (Eds.), *Artificial life XII: proceedings of the twelfth international conference on the simulation and synthesis of living systems* (pp. 687–693). Cambridge: MIT Press.

Schmickl, T., & Crailsheim, K. (2008). Trophallaxis within a robotic swarm: bio-inspired communication among robots in a swarm. *Autonomous Robots*, 25, 171–188.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379–423, 623–656.

Sperati, V., Trianni, V., & Nolfi, S. (2010). Evolution of self-organised path formation in a swarm of robots. In M. Dorigo, M. Birattari, G. A. Di Caro, R. Doursat, A. P. Engelbrecht, D. Floreano, L. M. Gambardella, R. Gross, E. Şahin, T. Stützle, & H. Sayama (Eds.), *Lecture notes in computer science: Vol. 6234. Proceedings of the 7th international conference on swarm intelligence (ANTS 2010)* (pp. 155–166). Berlin: Springer.

Stirling, T., Wischmann, S., & Floreano, D. (2010). Energy-efficient indoor search by swarms of simulated flying robots without global information. *Swarm Intelligence*, 4, 117–143.

Thrun, S. (2003). Robotic mapping: a survey. In G. Gerhard Lakemeyer & B. Nebel (Eds.), *Exploring artificial intelligence in the new millennium* (pp. 1–35). San Francisco: Morgan Kaufmann.

Thrun, S., & Liu, Y. (2005). Multi-robot SLAM with sparse extended information filters. In P. Dario & R. Chatila (Eds.), *Springer tracts in advanced robotics: Vol. 15. Robotics research. The eleventh international symposium* (pp. 254–266). Berlin: Springer.

Trianni, V. (2008). *Studies in computational intelligence: Vol. 108. Evolutionary swarm robotics. Evolving self-organising behaviours in groups of autonomous robots*. Berlin: Springer.

Trianni, V., & Nolfi, S. (2011). Engineering the evolution of self-organising behaviours in swarm robotics: A case study. Artificial Life, 17(3) (to appear).

Vaughan, R. T., Støy, K., Sukhatme, G. S., & Matarić, M. J. (2002). LOST localization-space trails for robot teams. *IEEE Transactions on Robotics and Automation*, 18(5), 796–812.

Vickerstaff, R. J., & Di Paolo, E. A. (2005). Evolving neural models of path integration. *Journal of Experimental Biology*, 208, 3349–3366.

Wehner, R. (2003). Desert ant navigation: how miniature brains solve complex tasks. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural and Behavioral Physiology*, 189(8), 579–588.

Werger, B., & Matarić, M. J. (1996). Robotic "food" chains: Externalization of state and program for minimal-agent foraging. In P. Maes, M. J. Matarić, J.-A. Meyer, J. Pollack, & S. W. Wilson (Eds.), *From animals to animats 4. Proceedings of the fourth international conference on simulation of adaptive behavior (SAB 96)* (pp. 625–634). Cambridge: MIT Press.

Zeil, J., Boeddeker, N., & Stürzl, W. (2009). Visual homing in insects and robots. In D. Floreano, J.-C. Zufferey, M. V. Srinivasan, & C. Ellington (Eds.), *Flying insects and robots* (pp. 87–100). Berlin: Springer.