Swarm Intell (2008) 2: 73–95
DOI 10.1007/s11721-008-0017-1

Evolving coordinated group behaviours through maximisation of mean mutual information

Valerio Sperati · Vito Trianni · Stefano Nolfi

Received: 26 November 2007 / Accepted: 5 August 2008 / Published online: 17 September 2008

Abstract A well known problem in the design of the control system for a swarm of robots concerns the definition of suitable individual rules that result in the desired coordinated behaviour. A possible solution to this problem is given by the automatic synthesis of the individual controllers through evolutionary or learning processes. These processes offer the possibility to freely search the space of the possible solutions for a given task, under the guidance of a user-defined utility function. Nonetheless, there exist no general principles to follow in the definition of such a utility function in order to reward coordinated group behaviours. As a consequence, task dependent functions must be devised each time a new coordination problem is under study. In this paper, we propose the use of measures developed in Information Theory as task-independent, implicit utility functions. We present two experiments in which three robots are trained to produce generic coordinated behaviours. Each robot is provided with rich sensory and motor apparatus, which can be exploited to explore the environment and to communicate with other robots. We show how coordinated behaviours can be synthesised through a simple evolutionary process. The only criteria used to evaluate the performance of the robotic group is the estimate of mutual information between the motor states of the robots.

Keywords Evolutionary robotics · Information theory · Mutual information

Electronic supplementary material The online version of this article (http://dx.doi.org/10.1007/s11721-008-0017-1) contains supplementary material, which is available to authorized users.

V. Sperati · V. Trianni (✉) · S. Nolfi
Laboratory of Autonomous Robotics and Artificial Life, Institute of Cognitive Sciences and Technologies, CNR, via S. Martino della Battaglia, 44-00185 Rome, Italy
e-mail: vito.trianni@istc.cnr.it

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

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

1 Introduction

The design of the control system for a swarm of robots is not a trivial enterprise. Above all, it is difficult to define the individual rules that produce a desired swarm behaviour without an a priori knowledge of the system features. For this reason, evolutionary or learning processes have been widely used to automatically synthesise group behaviours (see, for instance, Matarić 1997; Quinn et al. 2003; Baldassarre et al. 2007). In this paper, we investigate the use of information-theoretic concepts such as *entropy* and *mutual information* as utility functions for mobile robots, which adapt on the basis of an evolutionary or learning process. We believe that the use of implicit and general purpose utility functions—fitness functions or reward/error measures—can allow evolution or learning to explore their search space more freely, without being constrained by an explicit description of the desired solution. In this way, it is possible to discover behavioural and cognitive skills that play useful functionalities, and that might be hard to identify beforehand by the experimenter without an a priori knowledge of the system under study. Such task-independent utility functions can be conceived as universal intrinsic drives toward the development of useful behaviours in adaptive embodied agents. A second relevant aspect of this approach, which will be investigated in future work, concerns the combination of implicit and task-independent utility functions with explicit and task-dependent ones, in order to favour the development of behavioural and cognitive skills that serve specific requested functionalities, but that would be hard to develop through explicit descriptions only.

Relevant work in this direction has been carried out recently. Sporns and Lungarella (2006) demonstrated how the maximisation of the information structure of the sensory states experienced by embodied and situated agents might lead to the development of useful behavioural skills. The agent is a simulated arm provided with visual and tactile sensors, placed in an environment including an object that moves in a random direction at constant speed. The object is characterised by a uniform colour which can be distinguished from the randomly coloured pixels of the background. By evolving agents on the basis of the information structure of their experienced sensory states, the authors observed the development of useful behavioural skills consisting in the ability to foveate and to touch the moving object. Capdepuy et al. (2007) demonstrated how a wide range of coordinated collective behaviours can be developed by having a population of creatures situated in the same environment which adapt by maximising their *empowerment*—an individual-based utility function that measures the information transfer between the actions produced by the agent and the sensory states later experienced by the agent itself. The term “empowerment” refers to the fact that this measure encodes the perceived amount of influence or control that the agent has over the world. Prokopenko et al. (2006a) demonstrated how the maximisation of a measure of the spatio–temporal coordination of the body parts of a simulated snake-like robot can lead to useful coordinated behaviour. In particular, the authors showed how a utility function that minimises the irregularity over both space and time of the state of the actuators connecting the body parts of the robot can lead to the evolution of effective locomotion behaviours.

In this paper, we exploit information-theoretic measures to evolve collective behaviours for a homogeneous group of robots. In particular, we demonstrate how the use of a utility function that maximises the mutual information in state and time between the motor states of wheeled robots leads to the evolution of a variety of effective coordinated behaviours. In the present study, three robots driven by identical neural controllers prove capable of displaying behaviours that are both structured and coordinated. We define a “structured” behaviour as a temporal sequence of several elementary behaviours, where the latter are sequences of atomic actions that produce a well defined outcome (e.g., “move-straight”, “move-to-light”, "avoid-obstacle", etc.). For instance, an oscillatory behaviour in which a single robot moves back and forth from a light bulb is structured as it can be described as a periodic sequence of "move-to-light" and "move-away-from-light" behaviours. In contrast, sequences of random atomic actions would not be considered structured.

We present two sets of experiments which differ by the environmental cues available to the robots. In the first experiment, referred to as E_l, robots evolve within an arena presenting a clearly distinguishable cue, that is, a light bulb perceivable from every location. In the second experiment, referred to as E_d, there is no light bulb that provides environmental cues to be exploited, and the robots have to autonomously create the conditions required to perform structured and coordinated behaviours. We show how the proposed measure leads to the evolution of a rich—nontrivial—repertoire of coordinated behaviours. Moreover, the paper assesses the effectiveness of the proposed methodology through the use of realistic simulations and through the test of the solutions evolved in simulation on the physical robots.

The rest of the paper is organised as follows. In the next section, we briefly review the relevant aspects of information theory. In Sects. 3 and 4, we describe the experimental setup and the results obtained. In Sect. 5, we perform a scalability analysis to test the performance of the evolved controllers with large groups of robots. Finally, in Sect. 6, we discuss the main contributions of the paper and we draw our conclusions.

2 Short introduction to information theory

In this section, we briefly discuss the information theory concepts and measures first introduced by Shannon (1948), used in the definition of the task independent utility function described in Sect. 3.3. Regarding notation, we follow Feldman's style: we use capital letters to indicate random variables, and lowercase to indicate a particular value of a variable (Feldman 2002). For example, let X be a discrete random variable. The variable X may take on the values $x \in \mathcal{X}$. Here, $\mathcal{X}$ is the finite set of M possible values (or states) for X, referred to as the *alphabet* of X. The probability that X takes on the particular value x is written $p(X = x)$, or just $p(x)$ (first order probability density function). We may also form joint and conditional probabilities. Let Y be another random variable with $Y = y \in \mathcal{Y}$. The probability that $X = x$ and $Y = y$ is written $p(X = x, Y = y)$, or simply $p(xy)$ (second order probability density function), and is referred to as a joint probability. The conditional probability that $X = x$ given $Y = y$ is written $p(X = x|Y = y)$, or simply $p(x|y)$. Now, we can introduce the Shannon entropy equation which is formally defined as:

$$H[X] = -\sum_{x\in\mathcal{X}} p(x) \cdot \log_2 p(x). \tag{1}$$

The entropy $H[X]$ is equal to zero if the variable X always takes on the same value. The maximum value is equal to $\log_2 M$, and it is obtained when X takes on all M possible values of the *alphabet* $\mathcal{X}$ with the same probability ($p = \frac{1}{M}$). There are many interpretations of the meaning of Shannon entropy. In our case, we consider entropy as "a precise measure of the amount of freedom of choice in an object; an object with many possible states has high entropy" (see Prokopenko and Wang 2003). The same formula and interpretation is applicable to a joint distribution:

$$H[XY] = -\sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} p(xy) \cdot \log_2 p(xy). \tag{2}$$

Note that, by definition, $H[XY] \leq H[X] + H[Y]$. The equality is obtained if and only if X and Y are statistically independent. Given a conditional distribution we can define the conditional entropy:

$$H[X|Y] = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(xy)\cdot \log_2 p(x|y). \tag{3}$$

The conditional entropy quantifies the remaining entropy about X, given that the value of Y is known. Note that $H[X|Y]=0$ if and only if the value of X is completely determined by the value of Y. Conversely, $H[X|Y]=H[X]$ if and only if X and Y are statistically independent. It is quite useful to see that the equation of joint entropy can be reformulated in terms of marginal entropy and conditional entropy:

$$H[XY] = H[X] + H[Y|X] = H[Y] + H[X|Y]. \tag{4}$$

Finally, we present the Mutual Information (*MI*) which is formally defined as:

$$MI[X;Y] = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(xy)\cdot \log_2 \frac{p(x)\cdot p(y)}{p(xy)}. \tag{5}$$

The properties of *MI* are more evident if we re-write the above formula in terms of the marginal entropy and the joint entropy and in terms of the marginal entropy and the conditional entropy:

$$MI[X;Y] = H[X] + H[Y] - H[XY], \tag{6}$$

$$MI[X;Y] = H[X] - H[X|Y] = H[Y] - H[Y|X]. \tag{7}$$

Looking at (5), it is possible to notice that $MI[X;Y]=0$ if the two variables are statistically independent. On the other hand, (6) shows that $MI[X;Y]=0$ if the two variables have zero entropy. Finally, note that $MI[X;X]=H[X]$, and $MI[X;Y]=MI[Y;X]$.

The interpretation of *MI* is quite clear looking at (7). Feldman describes *MI* as "the reduction in uncertainty of one variable due to knowledge of another. If knowledge of Y reduces our uncertainty of X, then we say that Y carries information about X" (Feldman 2002). In other words, if X and Y are independent variables, the mutual information that one variable brings about the other is null. On the other extreme, mutual information is maximised if the knowledge of one variable is sufficient to completely describe the other variable. In practise, *MI* can be used as a powerful index of correlation: the greater the value of *MI*, the more correlated the two variables. The great advantage of *MI* is that it takes into account both linear and nonlinear dependencies (Lungarella and Pfeifer 2001).

The above measures—and other derivations—have been exploited successfully in many different fields. In ethology, information entropy has been used to describe the interplay between pheromone molecules and ants' movements. Observing foraging behaviour in ants, Van Dyke Parunak and Brueckner (2001) showed that an increase in entropy at the micro-level of the chemical particles is compatible with a reduction of disorder at the macro-level of the ants' movements. Brenner et al. (2000) used information entropy to describe the behaviour of the visual system of the fly. The authors showed how the fly's response to the environmental features is dynamically adapted in order to maximise the information inflow. In neurosciences, the dynamics observable in the human brain have been studied under the light of information theory (Tononi et al. 1994, 1996, 1998; Sporns et al. 2000). A new measure called *neural complexity* (C_N) captures some aspects of the interplay between the functional segregation of different cortical areas and their global integration during perception and behaviour. C_N is shown to be high when functional segregation coexists with global integration, and to be low when the components of a system are either completely independent (segregated) or completely dependent (integrated). In robotics, Olsson et al. (2005)

proved that the perceptions of a SONY AIBO robot can be treated in an efficient and computationally economic way if sensors can adapt to the statistical properties of the incoming signals. In another interesting work, entropy and mutual information have been applied to the sensory channels of a two wheeled simulated robot (Tarapore et al. 2004, 2006). These measures were used to classify different behaviours, such as exploring the environment, searching for red objects and tracking them. The authors argued that information theory can provide useful methods to discover the "fingerprints" of particular agent–environment interactions. Similarly, Lungarella and Pfeifer (2001) used entropy and mutual information to analyse the input data obtained by a robotic arm holding a colour camera. The authors compared coordinated movements (i.e., foveation on a red object), with uncoordinated ones (i.e., random movements), detecting clear informational structures in the first case. Comparable results were obtained by Lungarella and Sporns (2005) and Lungarella et al. (2005), using a robotic setup very similar to the previous. The authors argued that coordinated sensory–motor activity induces information structures in the sensory experience.

Conversely, it has been argued that sensory–motor coordination can be achieved by rewarding information structure in the sensory or motor experience of autonomous robots (Sporns and Lungarella 2006; Capdepuy et al. 2007; Klyubin et al. 2005a, 2005b; Prokopenko et al. 2006a). These works—as well as the work presented in this paper—belong to a novel methodology in evolutionary robotics called *information-driven evolution*, in which generic information theoretic criteria are exploited as utility functions. "The solutions obtained by *information-driven evolution* can be judged by their degree of approximation to the direct evolution results. A good approximation will indicate that the chosen criteria capture the information-theoretic core of selection pressures, leading to task-less adaptation" (Prokopenko et al. 2006b). In this work, we focus on applying mutual information as the generic utility function to obtain task-less adaptation, as described in the following section.

3 Experimental setup

As mentioned in the previous section, $MI[X;Y]$ can be seen as a powerful measure to grasp the correlation between two stochastic variables X and Y. Moreover, maximising *MI* also corresponds to maximising the entropy of the single variables $H[X]$ and $H[Y]$,[1] which is related to an higher probability of observing X or Y in multiple states. In this paper, we study whether *MI* can be used to evolve coordinated behaviours in a group of robots. However, the application of such a measure as utility function for an evolutionary robotics experiment is not straightforward. Given the experimental setup, it is necessary to define which are the stochastic variables under observation, discretise them in a suitable way and compute the desired utility functions. We chose to maximise the mutual information of the *motor states* observed in a group of autonomous robots (see Sect. 3.3). In particular, we focus on the wheels' speed, which characterises the robot movements in the environment. In this way, we aim at evaluating the quality of the individual and group behaviour, without any reference to the sensory pattern perceived by the robots.

The experimental setup involves three wheeled robots provided with a neural controller and different types of actuators and sensors (see Sect. 3.1). Robots are placed in a 1 m side square arena surrounded by walls. In the experiment E_l, a light bulb is placed in the centre of the arena. The intensity of the light decreases quadratically with the distance from the light bulb, but it is anyway perceivable by the robots from every location in the arena.

[1] This is true if the joint entropy is kept constant, see (6).

Therefore, the light bulb provides a clearly distinguishable environmental cue to be exploited by the robots for coordination. In the experiment E_d, such environmental cue is not present, making the coordination between the robots more difficult to achieve.

We performed 20 *evolutionary runs* per experiment, in order to establish the viability of the approach, varying the initial population of genotypes. Each evolutionary run is iterated for 200 generations, in which the population is evaluated and genotypes are selected for reproduction on the basis of an estimate of their fitness (see Sect. 3.2). This estimate is obtained by testing each genotype 10 times—i.e., we perform 10 independent *trials* randomly varying the initial conditions (see Sect. 3.3). The best evolved genotypes resulting from each evolutionary run is then selected for a qualitative and quantitative analysis, presented in Sect. 4.

3.1 The robot and the neural controller

The experiments presented in this paper are performed using the *e-puck* robots (see Fig. 1 left), which are wheeled robots with a cylindrical body having a diameter of 70 mm (Mondada and Bonani 2007; Cianci et al. 2007). A rich subset of the sensory–motor features of the *e-puck* has been exploited, as detailed below. In fact, by using an implicit and task-independent fitness function, we do not define a particular goal to be pursued by the robots. As a consequence, we do not know in advance the sensory–motor features that can be exploited to maximise the fitness function. We therefore decided to provide the robots with a rich set of sensors and actuators in order to let the evolutionary process free to explore a wide set of possible solutions.

Each robot is provided with various sensory systems that allow it to perceive the environment, including the other robots. Among these, we make use of infrared proximity sensors, ambient light sensors, and a VGA camera pointing in the direction of forward motion. Moreover, the robots can communicate with their neighbours in two different ways. They can light up the 8 red LEDs distributed around their body, in order to be detected by cameras of the other robots. Additionally, robots can exploit their wireless Bluetooth interface to send and receive short messages.

The robots are controlled by artificial neural networks, whose parameters are set by an evolutionary algorithm. A single genotype is used to create a group of robots with an identical control structure—a homogeneous group. Each robot is controlled by a fully connected two layer neural network, characterised by an input layer with leaky integrator neurons and

Fig. 1 *Left*: The e-puck robot developed at the Swiss Federal Institute of Technology in Lausanne (EPFL) (Mondada and Bonani 2007). *Right*: A close up view of the environment with the light bulb in the centre and three robots

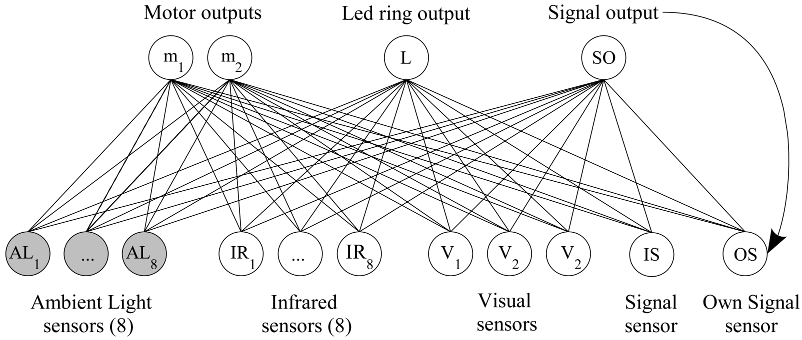

Fig. 2 The architecture of the neural controller. The grey neurons and the corresponding connections are used in the experiment E_l only, while the other neurons are common to both setups

by an output layer of motor neurons (see Fig. 2). The activation of the output neurons is computed as the weighted sum of all input units and the bias, filtered through a sigmoid function:

$$O_j(t) = \sigma\left(\sum_i w_{ij} I_i(t) + \beta_j\right), \qquad \sigma(z) = \frac{1}{1+e^{-z}}, \tag{8}$$

where $I_i(t)$ corresponds to the activation of the ith sensory neuron at time t, w_{ij} is the weight of the synaptic connection between the sensory neuron i and the motor neuron j, and β_j is a bias term. Sensory neurons are leaky integrators, that is, they hold a certain amount of their activation from the previous time step, and the effect of the previous state on their current state is determined by their time constant:

$$I_i(t) = \tau_i \cdot I_i(t-1) + (1-\tau_i) \cdot \mathcal{I}_i(t), \tag{9}$$

where τ_i is the time constant of the ith neuron, and $\mathcal{I}(t)$ is the sensory input at time t.

The activations of the output neurons are real valued numbers in the range [0.0, 1.0], and are used to control the actuators of the robot (see Fig. 2). Two motor neurons (m_1 and m_2) encode the desired speed of the two motors which control the two corresponding wheels. The activation of each neuron is linearly scaled in the range $[-2\pi, 2\pi]$ rad/sec, and used to set the desired angular speed of the corresponding motor. One motor neuron (L) controls the red LEDs: all eight LEDs are switched on or off depending on whether the activation of the motor neuron is above or below an arbitrary threshold of 0.9. Finally, one motor neuron (SO) encodes the value of the communication signal produced by the robot at each cycle, which varies in the range [0.0, 1.0]. This signal is transmitted to the other robots through the wireless Bluetooth interface.

Concerning the sensory inputs, they are set by the robot sensors after normalising their value onto the range [0.0, 1.0]. Eight sensory inputs are dedicated to the infrared proximity sensors ($IR_i, i = 1, \dots, 8$), which can detect an obstacle up to a distance of approximately 25 mm (see Fig. 2). Three sensory inputs ($V_i,\ i = 1, \dots, 3$) encode the presence of nearby robots—provided that they have their red LEDs switched on—as detected by the camera: the image that is grabbed at each cycle is pre-processed, in order to extract the percentage of pixels that have a predominant red colour within the following three vertical visual sectors: $[-18°, -6°]$, $[-6°, +6°]$, and $[+6°, +18°]$. The two remaining sensory inputs are dedicated to the communication signal: one input (IS) encodes the average signal produced by all the robots placed in the arena, the other input (OS) encodes the signal produced by the robot itself during the previous cycle. Additionally, in the experimental setup that includes the light bulb, the robots are provided with eight further sensory inputs, which are dedicated to the ambient light sensors (AL_i, $i = 1, \dots, 8$), shown in grey in Fig. 2.

In the experiments performed in simulation, the state of the infrared and ambient light sensors has been simulated through a sampling technique (Miglino et al. 1995). The visual sensors have been simulated through a ray tracing technique, by using 36 rays uniformly distributed over the camera range. All sensors have been subjected to noise implemented as a random value with a uniform distribution in the range $[-0.05, 0.05]$, added to the state of each simulated sensor. The use of simulated noise should favour the portability of the controllers evolved in simulation to the physical robots (see Jakobi 1997, for a detailed discussion about this topic).

3.2 The evolutionary process

The free parameters of the robot's neural controller are adapted through an evolutionary process (Nolfi and Floreano 2000). The initial population consists of 100 randomly generated binary genotypes, that encode the connection weights, the bias terms and the time constants of 100 corresponding neural controllers. Each parameter is encoded by 8 bits, and its value is linearly scaled from the range $[0, 255]$ to the range $[-5.0, 5.0]$ in the case of connection weights and bias terms, and in the range $[0.0, 0.95]$ in the case of time constants. The 20 best genotypes of each generation were allowed to reproduce by generating five copies each, with 4% of their bits replaced with a new randomly selected value, excluding one copy (elitism). The evolutionary process lasted 200 generations.

Each genotype is translated into 3 identical neural controllers which are downloaded onto three identical robots (i.e., the robots are homogeneous). Each team of 3 robots was tested for 10 trials, lasting 200 seconds (i.e., 2000 simulation cycles of 100 ms each). The performance of the genotype is the average fitness, as computed by (12), over 10 trials. At the beginning of each trial, the three robots are placed in the arena with a random position and orientation. In case of collision the team is repositioned randomly again. The evolutionary process has been conducted in simulation.[2] The best evolved neural controllers have been tested with physical robots.

3.3 The fitness function

Evolving individuals are selected on the basis of a fitness function which measures the Mutual Information *MI* between the motor states X_i of all possible robot pairs. The maximisation of *MI* should drive evolution towards the development of coordinated behaviours. In fact, high values of *MI* correspond to motor states that are positively correlated: the knowledge of motor state X_i gives information about motor state X_j and vice versa. In other words, X_i and X_j result from processes that we can describe as "coordinated". Moreover, since the maximisation of *MI* also requires the maximisation of the entropy of the motor state X_i of each robot, this fitness function rewards evolving robots for the ability to produce structured behaviours. In fact, entropy is maximised not only by very random behaviours, but also by very structured behaviours that systematically vary through time. In particular, periodic sequences of equally frequent elementary behaviours such as "move-forward", "move-backward", "turn-left", and "turn-right" allow the robot to uniformly cover many possible motor states, therefore maximising entropy.

[2] Using a similar setup, a single evolutionary run—i.e., 200 generations, 100 individuals, 10 trials per individual, 200 seconds per trial—performed with physical robots would last longer than one year.

For the purpose of computing the fitness function as the *MI* between the motor states of a robot pair, we need to define a discrete variable X that accounts for the current motor state—the wheels' speed. To avoid that motor state variations are caused by the random noise injected in the simulation, we filter the motor state through a slow moving average. In this way, for robots not having internal dynamics, systematic variations of X_i can solely be produced by exploiting the dynamics of the robot/environment interaction (i.e., by exploiting sensory–motor coordination). The activation values m_j, $j = 1, 2$, of the two motor neurons controlling the wheels has been averaged through time into the variables M_j:

$$M_j(t) = \gamma \cdot M_j(t-1) + (1-\gamma) \cdot m_j(t), \quad j = 1, 2, \tag{10}$$

where $m_j(t) \in [0.0, 1.0]$ indicates the current activation of the motor neuron j and $\gamma = 0.9$ is a fixed time constant that represents the rate at which $M_j(t) \in [0.0, 1.0]$ changes over time. This moving average also channels the evolutionary process towards the synthesis of behaviours that extend for sensible time durations.[3] The overall motor state X of a robot is a discrete variable computed according to the following equation:

$$X = \lfloor M_1 \cdot 5 \rfloor + \lfloor M_2 \cdot 5 \rfloor \cdot 5, \tag{11}$$

where $\lfloor M_j \cdot 5 \rfloor$ means that the value M_j has been discretised into the integer range $[0, 4]$, encoding all possible activation values of the motor neuron into five discrete states.[4] As a consequence, X takes on integer values in the range $[0, 24]$.

In order to compute the *MI* of a robot pair, the value X_i of each robot $i = 1, 2, 3$ is recorded in every cycle, obtaining statistics about the states encountered during a trial. On the basis of these statistics, it is possible to estimate the probability distribution $p(X_i = x)$ and the joint distribution $P(X_i = x, X_j = y)$ needed to compute $MI[X_i; X_j]$, according to Equation (5). Having estimated the probability distribution, the fitness function F of the group of robots in a trial is calculated on the basis of the following equation:

$$F = \binom{N}{2}^{-1} \sum_{i=1}^{N} \sum_{j=i+1}^{N} MI[X_i; X_j] \cdot \frac{20-c}{20}, \tag{12}$$

where N is the number of robots, and c is the number of times in which one of the robots collided against a wall or against another robot, truncated to 20. Notice that the sum is extended to all possible robot pairs, and it is normalised by the total number of different pairs with N robots. The second term of the fitness function has been introduced in order to reward robots for the ability to avoid collisions. All robots are randomly repositioned whenever a collision is detected: in this way, we bypass the problem of accurately simulating the physical interactions during a collision, offering the robots further possibilities to coordinate. Moreover, a maximum of 20 collisions per trial is allowed before the trial is stopped. These choices channel the evolution of good collision avoidance behaviours.

[3] Preliminary experiments conducted without the moving average produced behaviours that were coordinated, but neither periodic nor structured (result not shown). In these experiments, we observed that the motor state of each robot varied in a quasi-random way (e.g., alternating at each time-step very different actions such as move-forward, move-backward, turn-right, turn-left); therefore, maximising the individual entropy without actually being structured or periodic. Such variations were produced by achieving and maintaining a given relative position with respect to an obstacle or to the other robots, so that each movement resulted in a large variation of the sensory pattern.

[4] The activation value equal to $M_j = 1.0$ is considered as state 4.

The maximum value of F is obtained when no collisions are detected and all robot pairs have maximum MI. Since X can assume 25 different states, the fitness takes values in the range $[0.0, \log_2 25]$. It is worth noting, however, that the maximum value cannot be achieved in practise. The main reason for this is that the individual entropy cannot be maximised because robots are embodied and their dynamical interaction with the environment—as it is defined by the neural controller—constrains the number of motor states visited during the robot's lifetime, and their relative frequency. Moreover, the motor state X is the result of a moving average with a fixed time constant, which influences X's variability. Finally, the computation of the MI includes the initial transitory phase during which the robots try to achieve a coordinated behaviour.

4 Results

In this section, we describe the results obtained in the two experiments E_l and E_d. As detailed in the following sections, in both experiments the evolved robots display behaviours that are structured (i.e., they consist of a sequence of atomic movements with varying time durations), periodic (i.e., the sequence of atomic movements is repeated through time), and coordinated (i.e., the different robots tend to produce the same structured behaviour in a synchronised manner). From a qualitative point of view, the evolved behaviours vary considerably between the two experiments, and also across the different evolutionary runs of the same experiment.

4.1 Experiment E_l

In the experiment E_l, the robots are situated in a 1 m side square arena presenting a light bulb, which can be perceived by means of the robot's ambient light sensors. As mentioned above, we performed 20 evolutionary runs, each time starting with a different randomly generated population. After the evolutionary process, we selected the best individual of each run for post-evaluation. In this case, the fitness of each individual was further evaluated for 500 trials, using (12). The results obtained are summarised in Table 1, in which we show mean and standard deviation over the 500 trials of the fitness F, of the mean mutual information $\widehat{MI}$ over all possible robot pairs, and of the mean entropy $\widehat{H}$ computed over all robots. The results of the post-evaluation show that the average fitness varies between 1.70 and 3.24, obtained, respectively, in runs 1 and 16. Given that F has been explicitly constructed as a task independent and implicit utility function, the absolute value of F is not very informative about the quality of the evolved behaviour. Recall that the absolute value of F is mainly given by the $\widehat{MI}$. The latter is constrained by the robots' embodiment which limits the number of possible motor states actually visited during the robot's lifetime. A qualitative analysis revealed that 18 out of 20 evolutionary runs resulted in controllers that produce structured and coordinated behaviours (see the runs indicated by a black dot in Table 1). This is a first result proving that the proposed methodology is viable: mutual information can be exploited as a generic utility function to obtain task-less adaptation in a group of robots.

4.1.1 Behavioural analysis

The qualitative inspection of the results obtained indicates that the robots always display structured and coordinated behaviours. Generally, the environmental cue offered by the

Table 1 Experiment E_l: fitness F, mean mutual information $\widehat{MI}$ and mean entropy $\widehat{H}$ computed by testing in simulation the best evolved controller of each evolutionary run for 500 trials of 2000 cycles. Mean value and standard deviation are shown. The symbol • indicates a run in which the best evolved individuals clearly show behaviours that an external observer can judge as structured and coordinated

Run	F	$\widehat{MI}$	$\widehat{H}$
1	1.70 ± 0.27	1.72 ± 0.28	2.55 ± 0.39
2 •	2.81 ± 0.14	2.84 ± 0.11	3.55 ± 0.05
3 •	1.91 ± 0.20	1.93 ± 0.17	2.97 ± 0.08
4 •	2.99 ± 0.21	3.02 ± 0.18	3.96 ± 0.04
5 •	2.97 ± 0.13	2.99 ± 0.11	3.84 ± 0.08
6 •	2.50 ± 0.07	2.50 ± 0.07	3.24 ± 0.13
7 •	2.41 ± 0.14	2.42 ± 0.12	3.26 ± 0.15
8 •	2.19 ± 0.19	2.24 ± 0.16	3.43 ± 0.08
9 •	2.40 ± 0.18	2.43 ± 0.14	3.32 ± 0.05
10 •	2.17 ± 0.20	2.18 ± 0.19	3.07 ± 0.13
11 •	2.73 ± 0.17	2.75 ± 0.13	3.51 ± 0.09
12 •	2.27 ± 0.15	2.29 ± 0.13	3.52 ± 0.19
13 •	2.38 ± 0.22	2.39 ± 0.21	3.19 ± 0.25
14 •	2.72 ± 0.13	2.75 ± 0.09	3.55 ± 0.06
15 •	2.47 ± 0.18	2.51 ± 0.12	3.23 ± 0.04
16 •	3.24 ± 0.14	3.25 ± 0.12	4.01 ± 0.06
17 •	2.49 ± 0.16	2.50 ± 0.13	3.55 ± 0.02
18	1.72 ± 0.17	1.75 ± 0.16	3.05 ± 0.21
19 •	2.99 ± 0.17	3.01 ± 0.14	3.96 ± 0.04
20 •	3.12 ± 0.17	3.14 ± 0.14	4.08 ± 0.06

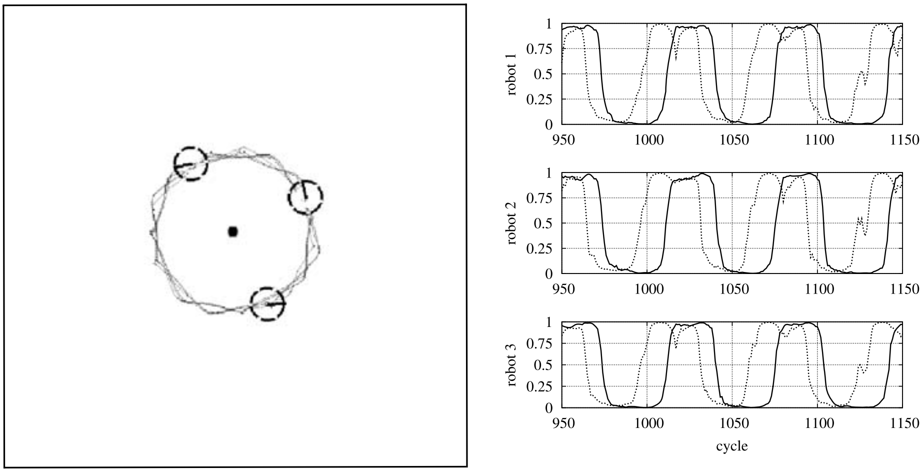

Fig. 3 Analysis of the behaviour produced by the best evolved controller in run 16 of experiment E_l. *Left*: trajectories of the robots. *Right*: activation of the motor neurons of each robot, plotted from cycle 950 to cycle 1150 to highlight the periodic motion of the robots. The *solid* and *dotted lines* indicate, respectively, the left and right motor neurons

light bulb is exploited by the robots to achieve the same relative position and to display a periodic, structured behaviour. Moreover, robots perform a coordinated behaviour through the synchronisation of their movements. Synchronisation is generally achieved through the exploitation of the communication signal only. Infrared sensors are generally exploited to avoid collisions with walls and with other robots, while visual information is often ignored.

A particularly interesting example of structured and coordinated behaviour is produced by the controller evolved in run 16 characterised by the highest mean performance (see Table 1). In this case, robots circle anticlockwise around the light bulb maintaining a distance of about 20 cm (see the trajectories of the robots shown in Fig. 3 and VideoEl-Run16 in the online supplementary material). While circling around the light bulb, robots display a structured behaviour composed of four atomic movements: (i) forward motion on the circle,

(ii) clockwise turn on the spot, (iii) backward motion on the circle, and (iv) anticlockwise turn on the spot. These atomic movements can be clearly identified looking at the plots in Fig. 3 right, in which we show the activation of the motor neurons that control the two wheels. Recall that maximum forward rotation corresponds to 1, while maximum backward rotation corresponds to 0. Starting at cycle 950, both wheels present forward rotation, resulting in forward movement on the circle. Afterwards, the activation of the right motor neuron sharply decreases to 0, leading to a clockwise rotation on the spot. Then, also the left motor activations drops to 0, resulting in backward motion. Finally, the right motor activation increases to 1, producing an anticlockwise rotation on the spot. After this, the robot starts again with forward motion.[5]

The above description accounts for the structure of the evolved behaviour. The coordination between the robots can be appreciated by observing how the motor activations of the three robots coincide in time (see Fig. 3 right). In short, robots are synchronised as they perform the same movements at the same time. The mechanism that the robots exploit to achieve and maintain synchronisation is based on communication, and on the fact that robots are homogeneous. An individual robot mainly signals during forward motion, and stops signalling as soon as the clockwise movement starts. All robots perform the same individual movements, which synchronise on the basis of the mutual interactions through communication. If an external signal is perceived, the robot keeps moving forward until signalling stops. As a consequence, the clockwise movement cannot start until all robots are performing forward motion. When this happens, synchronisation is achieved. This simple mechanism—already observed by Trianni and Nolfi (2007)—is based on a simple reaction to the perception of a signal, that allows a robot to achieve and maintain a certain sensory–motor condition—referred to as *reset configuration* by Trianni and Nolfi (2007)—waiting for the other robots. Synchronised movements start when all robots achieve the reset configuration.

Having described qualitatively the evolved behaviour, the questions remain: how did this behaviour evolve? In what way is *MI* maximised? To answer these questions, it is necessary to observe the motor states X_i and to analyse their statistics. Figure 4 shows how the motor states vary through time. First of all, it is possible to notice how the initial coordination phase is followed by a phase in which the group behaviour is perfectly synchronised. Moreover, it is possible to observe how, during the coordinated phase, the motor states takes on many different values. In other words, the motor states of the robots vary considerably through

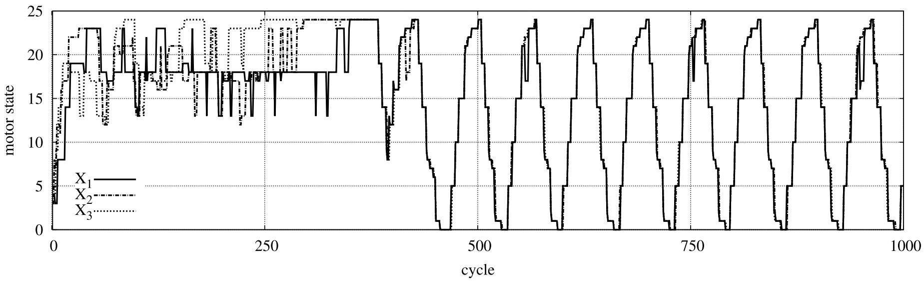

Fig. 4 The motor states of the three robots—computed using (11)—are plotted against the number of cycles. Notice the initial coordination phase, followed by synchronised movements

[5] See also http://laral.istc.cnr.it/esm/sperati-etal-si2008.html for videos and other supplementary material.

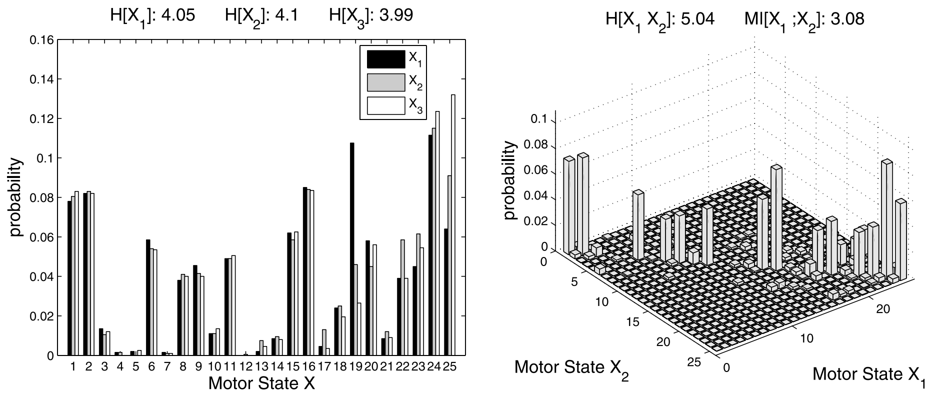

Fig. 5 *Left*: probability distribution for the motor states X_i of each robot $i = 1, 2, 3$. *Right*: probability distribution of the joint state $\langle X_1, X_2 \rangle$

time, which corresponds to a high individual entropy. Besides, once robots are synchronised, the motor states are highly correlated. This means that the joint entropy is minimised and the mutual information maximised.

Similar conclusions can be drawn looking at Fig. 5. In the left part, the histograms represent the probability $p(X_i = x), x \in [0, 24]$ estimated on a single trial. It is possible to notice how X_i takes on many different values with relatively high probability. As a consequence, the individual entropy $H[X_i]$ is rather high (see the individual values shown above the plot). Similarly, in the right part of Fig. 5, the 3D histogram represents the probability $p(X_1 = x_1, X_2 = x_2)$ estimated on the same trial.[6] Here, it is worth noting that the joint distribution takes values mainly on the diagonal $X_1 = X_2$, meaning that the probability of having $X_1 \neq X_2$ is rather low. As a consequence, we observe a small value for the joint entropy $H[X_1 X_2]$, and a high value for the mutual information $MI[X_1; X_2]$.

Owing to the above analysis, it is possible to claim that (i) structured behaviours maximise the individual entropy, because they are characterised by motor states that have sensible time duration and vary systematically across the range of possible values; (ii) coordinated behaviours maximise the mutual information, because they ensure that a certain motor state of one robot is correlated with the motor state of other robots; (iii) the homogeneity of the robots results in synchronisation behaviours that ensure the one-to-one correspondence of the motor states between robots.

4.1.2 Porting to reality

By testing with physical robots all controllers that proved successful in simulation, we observed qualitatively similar behaviours with respect to simulation in the majority of the evolutionary runs (12 out of 18 runs).[7] In all other cases, we observed a fairly good correspondence with simulation for individual behaviours, but not for coordination among robots. In fact, coordination was difficult to achieve and to maintain throughout a whole trial.

[6] The histograms for the other pairs $\langle X_1, X_3 \rangle$ and $\langle X_2, X_3 \rangle$ are extremely similar and have been omitted for space reasons.

[7] See VideoEl-Run16 in the online supplementary material.

Table 2 Experiment E_I: mean mutual information ($\widehat{MI}$) and mean entropy ($\widehat{H}$) computed by testing the evolved controllers on physical robots for 5 trials of 2000 cycles each. Here we show only the evolutionary runs that successfully transfer to reality from a qualitative standpoint. The column labelled 'ratio' indicates the ratio between the performance observed in hardware with respect to the performance observed in simulation

Run	$\widehat{MI}$	$\widehat{H}$	Ratio
2	2.29 ± 0.07	3.55 ± 0.03	0.81
3	1.34 ± 0.24	3.23 ± 0.16	0.70
5	2.62 ± 0.17	3.87 ± 0.05	0.88
6	2.24 ± 0.04	3.14 ± 0.09	0.90
9	2.22 ± 0.12	3.44 ± 0.05	0.92
11	1.93 ± 0.05	3.31 ± 0.07	0.71
12	1.83 ± 0.12	3.65 ± 0.22	0.81
13	1.78 ± 0.51	2.89 ± 0.28	0.75
16	2.82 ± 0.05	3.96 ± 0.01	0.87
17	1.89 ± 0.13	3.45 ± 0.06	0.76
19	2.42 ± 0.17	3.54 ± 0.10	0.81
20	2.55 ± 0.12	4.17 ± 0.05	0.82

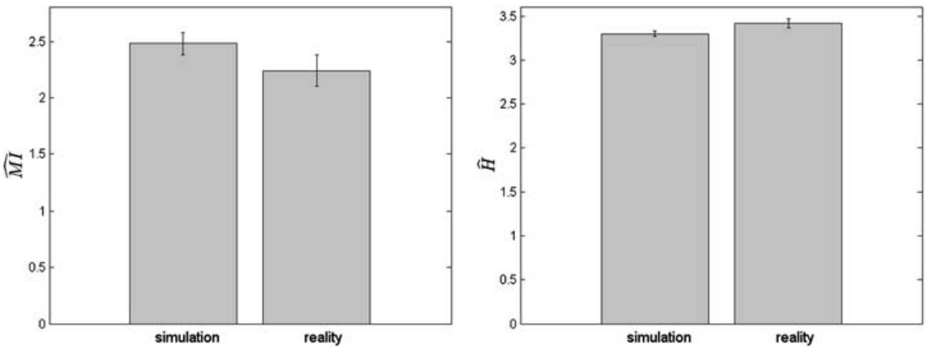

Fig. 6 Mean mutual information ($\widehat{MI}$) and entropy ($\widehat{H}$) computed by testing the best evolved controller of run 9 of experiment E_I in simulation and in reality for 20 trials of 2000 cycles. During the tests in hardware, the robots were situated in the same randomly generated positions and orientations that were used for the tests in simulation

In order to quantitatively determine the correspondence between tests with simulated and physical robots, we tested the evolved controllers by placing three real robots in locations randomly chosen from a set of 32 possible initial positions and 8 possible rotations. We performed 5 trials for each evolutionary run, and we measured the mean mutual information computed among all possible robot pairs. The results obtained are shown in Table 2, along with the ratio with the mean mutual information resulting from simulation. It is worth noting that the ratio between the mutual information observed in simulation and in the real environment is generally quite high, indicating that the behaviours tested in reality fairly correspond to those observed in simulation.

After this preliminary test performed on all evolutionary runs, we analysed in detail the best individual of run 9 (i.e., the individual with the highest ratio between the performance observed in simulation and in reality). We performed 20 further evaluations keeping exactly the same initial conditions in both simulated and real tests. We observed a good correspondence between the mean mutual information observed in simulation and in reality, as shown in Fig. 6 left. Similarly, the mean entropy over 20 trials computed on the tests with physical robots corresponds to the value obtained in simulation (see Fig. 6 right).

Table 3 Experiment E_d: fitness F, mean mutual information $\widehat{MI}$ and mean entropy $\widehat{H}$ computed by testing in simulation the best evolved controller of each evolutionary run for 500 trials of 2000 cycles. Mean value and standard deviation are shown. The symbol • indicates the runs in which the best evolved individuals display structured and coordinated behaviours. The symbol ◦ indicates the runs characterised by behaviours that degenerate with time

Run	F	$\widehat{MI}$	$\widehat{H}$	Run	F	$\widehat{MI}$	$\widehat{H}$
1 •	2.56 ± 0.15	2.57 ± 0.13	3.42 ± 0.03	11	1.45 ± 0.27	1.48 ± 0.25	2.88 ± 0.26
2 •	2.66 ± 0.12	2.71 ± 0.06	3.35 ± 0.08	12	1.46 ± 0.33	1.49 ± 0.33	2.22 ± 0.51
3 •	1.75 ± 0.15	1.77 ± 0.14	2.53 ± 0.13	13 •	1.85 ± 0.08	1.88 ± 0.07	2.99 ± 0.09
4	1.82 ± 0.13	1.84 ± 0.12	3.44 ± 0.25	14	1.24 ± 0.22	1.31 ± 0.21	2.72 ± 0.39
5 •	1.98 ± 0.11	1.99 ± 0.10	3.16 ± 0.06	15 •	2.59 ± 0.12	2.61 ± 0.09	3.31 ± 0.04
6 •	2.69 ± 0.15	2.72 ± 0.11	3.55 ± 0.04	16	1.42 ± 0.12	1.42 ± 0.12	2.35 ± 0.23
7	1.54 ± 0.07	1.54 ± 0.07	1.76 ± 0.06	17 ◦	2.22 ± 0.15	2.22 ± 0.14	2.55 ± 0.12
8 •	1.92 ± 0.14	1.94 ± 0.12	2.62 ± 0.12	18	1.27 ± 0.28	1.27 ± 0.28	2.07 ± 0.39
9 •	2.17 ± 0.12	2.19 ± 0.10	3.18 ± 0.19	19 ◦	1.93 ± 0.28	1.94 ± 0.28	2.31 ± 0.21
10 •	2.93 ± 0.07	2.95 ± 0.04	3.54 ± 0.04	20 •	2.02 ± 0.08	2.03 ± 0.08	2.72 ± 0.07

4.2 Experiment E_d

In the second set of experiments, the robots are situated in an arena without a light bulb. Moreover, robots are not provided with ambient light sensors. Also in this case, we performed 20 evolutionary runs, we selected the best individual of each run and we post-evaluated it in 500 different trials. As shown in Table 3, the evolved controllers present lower fitness values compared to the results obtained in experiment E_l. In this case, in fact, the fitness varies between 1.24 and 2.93, obtained in run 14 and 10, respectively. The qualitative analysis revealed that 11 out of 20 evolutionary runs converge toward structured and coordinated behaviours. In other two cases—namely runs 17 and 19—the average performance is rather high but robots display behaviours that are structured and coordinated only initially, and later degenerate toward nonstructured behaviours.

Despite the lower number of successful evolutionary runs, the proposed methodology for the evolution of coordinated behaviour still proves capable of producing good results in the majority of the tests performed (11 out of 20 evolutionary runs). The smaller number of successful evolutionary runs and the lower performance obtained in the average within experiment E_d is a consequence of the absence of the environmental cue that characterises experiment E_l. Indeed, all evolutionary runs of experiment E_l exploit such environmental cue, which gives a reference point that can be perceived from far away and that can be used by the robots to initiate and maintain a structured and coordinated behaviour. In contrast, the absence of the environmental cue forces the robots to search for other regularities that can be exploited for coordination. Given that the environment does not offer such obvious regularities, they must be extracted from the sensory-motor experience of the robots interacting with the *social environment*. Clearly, solutions of this kind are more difficult to evolve, because they are based on dynamical interactions among robots. However, as we show in the next section, a number of possible strategies exist to solve this problem.

4.2.1 Behavioural analysis

As mentioned before, the qualitative inspection of the evolved controllers allowed us to identify 11 evolutionary runs that produce structured and coordinated behaviours. Also in

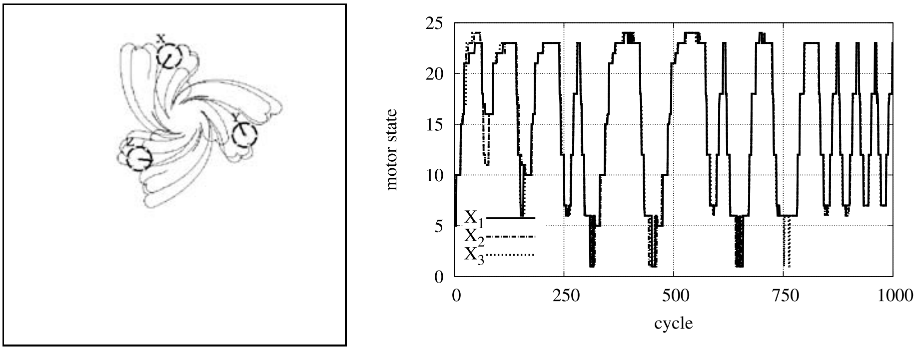

Fig. 7 *Left*: trajectories of the robots produced by the best evolved controller in run 6 of experiment E_d. *Right*: the motor states of the three robots are plotted against the number of cycles

this case, after an initial transitory phase, robots perform synchronised movements. Communication is exploited to achieve and maintain synchrony. The behaviours produced by the evolved controllers can be grouped into three strategies, described as follows.[8]

The first strategy—the most common one—encompasses the controllers evolved in runs 1, 3, 5, 6, 9, 15, and 20. An interesting example of this strategy is given by run 6, which presents the highest average fitness within its group. This strategy is characterised by robots that periodically aggregate and disband, performing oscillatory movements around the centre of mass of the group and faraway from the walls (see the trajectories in Fig. 7 left and VideoEd-Run6-9 in the online supplementary material). To do so, robots exploit vision, infrared proximity sensors and communication. Vision is mainly exploited in the aggregation phase during which robots get close one to the other assuming a triangular formation. When robots are close enough to perceive each other through the infrared proximity sensors, they disband moving backward. Due to relative differences in robots positions and orientations with respect to the centre of mass of the group, the behaviour of the three robots is not well coordinated during the first oscillatory movements. However, the robots quickly converge toward a well coordinated behaviour, as is apparent looking at the motor states plotted in Fig. 7 right. Notice also how the motor states vary through time, taking on many different values: this corresponds to a very structured behaviour, which is also well coordinated as the robots perform the same actions at the same time. Moreover, the oscillations have different amplitude and duration during a same trial, as can be noticed in Fig. 7 right. This fact indicates that robots are able to perform a variety of atomic movements, which can be triggered depending on the particular contingency the robots experience. Nevertheless, they prove capable of maintaining coordination even when switching between different oscillation modalities.

The second strategy encompasses the controllers evolved in runs 2, 10, and 13. The highest average fitness within this group is obtained by run 10 (see Fig. 8 and VideoEd-Run10 in the online supplementary material). In this case, robots do not interact visually or through their proximity sensors. They mainly produce a behaviour structured in a sequence of atomic movements, such as backward motion on a large circle followed by forward motion on a small circle. These movements are performed without any reference to the position and orientation of the other robots or to the position and orientation of the robot in the arena,

[8] See http://laral.istc.cnr.it/esm/sperati-etal-si2008.html.

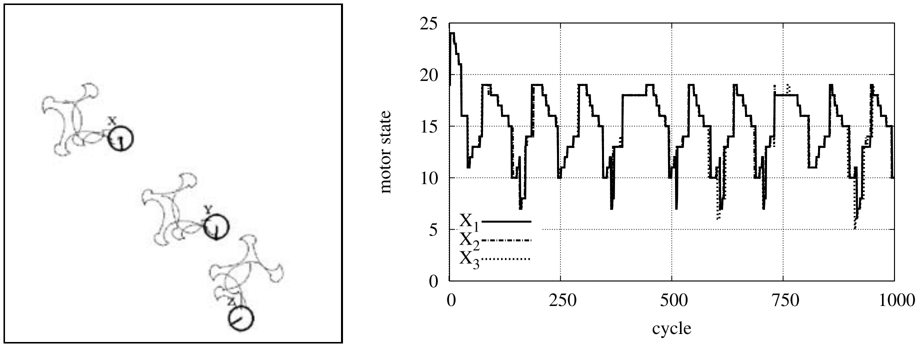

Fig. 8 *Left*: trajectories of the robots produced by the best evolved controller in run 10 of experiment E_d. *Right*: the motor states of the three robots are plotted against the number of cycles

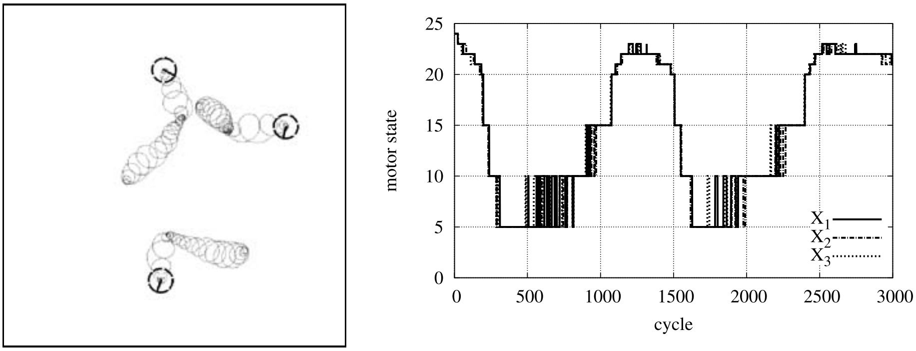

Fig. 9 *Left*: trajectories of the robots produced by the best evolved controller in run 8 of experiment E_d. *Right*: the motor states of the three robots are plotted against the number of cycles

provided that robots are located far enough from walls. Robots exploit only the communication signal to coordinate, and the robots display synchronised movements without keeping any relation between their relative positions in the arena. As a consequence, coordinated movements are performed since the very beginning of the trial, because there is no need to achieve a particular spatial formation (see the motor states plotted in Fig. 8 right).

Finally, the last strategy includes only the controller evolved in run 8 (see Fig. 9 and VideoEd-Run8 in the online supplementary material). This controller produces a peculiar behaviour characterised by four atomic movements that last from 10 to 40 seconds—i.e., a time span considerably longer than those observed in other evolutionary runs, which can be appreciated by looking at the motor states in Fig. 9 right—which are periodically repeated: (i) rotating several times to produce a nearly circular trajectory with a diameter of about 8 cm, (ii) rotating several times to produce a spiral trajectory with a diameter decreasing to 0 cm, (iii) rotating several times on the spot at full speed, (iv) rotating several times to produce a spiral trajectory with a diameter increasing from about 0 to about 8 cm. Also in this case, the movements of the robot are performed without any reference to the position and orientation of the other robots. However, we observed that visual information is exploited to switch between different rotating modes. Synchronisation of movements also characterises

Table 4 Experiment E_d: mean mutual information ($\widehat{MI}$) and mean entropy ($\widehat{H}$) computed by testing the evolved controllers on physical robots for 5 trials of 2000 cycles each. We show here only the evolutionary runs that successfully transfer to reality from a qualitative standpoint. The column labelled 'ratio' indicates the ratio between the performance observed in hardware and in simulation

Run	$\widehat{MI}$	$\widehat{H}$	Ratio
2	2.69 ± 0.06	3.36 ± 0.07	0.99
8	1.91 ± 0.03	2.65 ± 0.06	0.94
9	2.06 ± 0.10	3.21 ± 0.26	0.95
10	2.88 ± 0.05	3.51 ± 0.06	0.98
13	1.66 ± 0.06	2.90 ± 0.20	0.92

this behaviour (see the coordinated motor states in Fig. 9 right), and it is achieved and maintained exploiting communication only.

4.2.2 Porting to reality

By testing with physical robots all the controllers that proved successful in simulation, we observed good generalisation only in 5 out of 11 cases, namely runs 2, 8, 9, 10, 13 (see VideoEd-Run6-9, Run10 and Run8 in the online supplementary material). The main reason to explain the limited generalisation ability of these controllers is likely to be found in the fine grained interactions between robots that take place by means of the infrared proximity sensors. We found that proximity sensors differ significantly in sensitivity and perceptual range among different physical robots. Similar inter-robot differences were not systematically simulated, reducing the portability in hardware of the results obtained in simulation. Indeed, the evolutionary runs that produce qualitatively similar behaviour in simulation and in reality are characterised by limited interactions through infrared sensors.

For all evolutionary runs that properly generalise to the physical setup, the comparison of the mean mutual information $\widehat{MI}$ and mean entropy $\widehat{H}$ measured in simulation and in reality reveals a very good correspondence, as indicated by the high values of the ratio between the measures in the two conditions (see Table 4).

5 Scalability analysis

One important aspect to be investigated concerns the appropriateness of information theoretic measures for the synthesis of behaviours that scale up to larger groups. Such investigation can be performed either by running new evolutionary experiments involving more robots, or by testing whether the solutions evolved for small groups of robots properly scale when tested with larger groups. In this section, we describe the results obtained by following the latter approach, that is, by testing the controllers evolved with groups of 3 individuals—as described in the previous section—increasing the group size to 6, 12, 27, and 48 robots.

We first analysed the behaviour of the best evolved controllers of the successful evolutionary runs of experiment E_l. We increased the number of robots present in the environment, and we proportionally scaled up the arena size and the number of light sources present in the arena, in order to maintain a constant robot density and a uniform distribution of light sources. We performed experiments with 6, 12, 27, and 48 robots (see Table 5 for the corresponding experimental setup), and we recorded the fitness of the group computed with

Table 5 Parameters for the scalability analysis. We proportionally varied the number of robots, the arena size and the number of light bulbs, in order to maintain a constant density of robots and a uniform distribution of light sources

# robots	# light sources	arena size
6	2	2×1 m
12	4	2×2 m
27	9	3×3 m
48	16	4×4 m

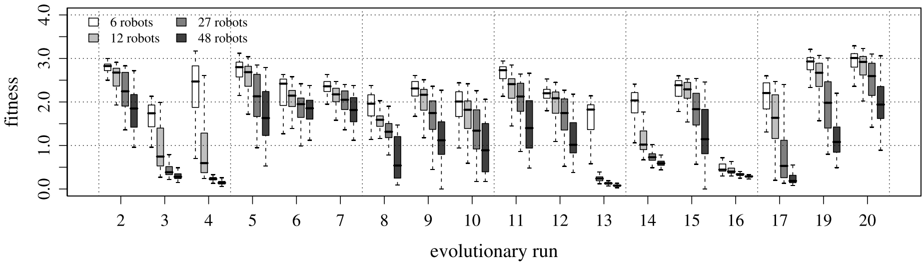

Fig. 10 Scalability analysis for experiment E_l. The *boxplot* shows, for each evolved controller, the performance obtained in tests with 6, 12, 27, and 48 robots. Each *box* represents the inter-quartile range of the data, while *the black horizontal line inside the box* marks the median value. The *whiskers* extend to the most extreme data points within 1.5 times the inter-quartile range from the *box*. Outliers are not shown

(12) over 100 trials. The results obtained are summarised in Fig. 10. It is possible to notice that most of the evolved controllers present good scalability, with a slight decrease in performance for increasing group size, due mainly to the slower convergence to synchronous movements for large groups. In general, robots are attracted by the closest light source, and therefore they distribute forming small groups around each light. Once aggregated, robots perform their structured periodic behaviour. Moreover, robots synchronise their movements with all the other robots, including those aggregated around a different light bulb. In fact, synchronisation is mostly achieved through the exploitation of the global communication channel, so that all robots in the arena can perform the same movements at the same time. In few cases, namely evolutionary runs 4, 13, 14, and 16, the evolved behaviour does not properly scale with increasing group size. In these cases, in fact, robots also exploit visual information about the position of teammates. As a consequence, they tend to aggregate with all other robots, forming large groups. In these conditions, either collisions are more frequent, or the local robot density is too high to allow coordinated movements.

In the case of experiment E_d, the scalability analysis was performed with the same modalities as for experiment E_l. The results obtained—presented in Fig. 11—suggest that scalability varies significantly depending on the behavioural strategy. In particular, the exploitation of visual information poses severe limitations to those controllers that strongly rely on it. All behaviours belonging to the first strategy described in Sect. 4.2.1 rely on visual information to aggregate in one place and to sustain their periodic behaviour. With increasing group size, robots tend to form large aggregates and become incapable of contemporaneously avoiding collisions and producing structured and coordinated behaviours. Similarly, the controller evolved in run 8 exploits visual information to produce its periodic behaviour. We observed that the stronger stimulation of the visual sensors prevent the robots to produce a structured behaviour. The only behaviours that properly scale are the ones evolved in run 2 and 10, both belonging to the second strategy described in Sect. 4.2.1. In fact, these behaviours do not rely on visual information, and synchronisation is achieved

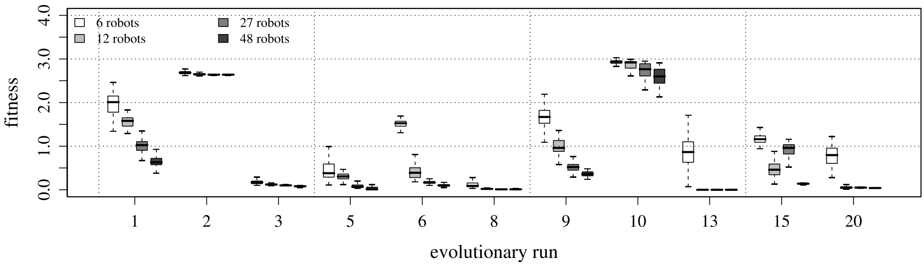

Fig. 11 Scalability analysis for experiment E_d. See Fig. 10 for more details

exploiting communication only.[9] Increasing the group size does not alter the ability of robots to immediately synchronise, so that performance is only slightly affected even with 48 robots, as can be appreciated by looking at Fig. 11.

To conclude, the scalability analysis we presented in this section proves that evolution of group behaviours exploiting information theoretic measures such as mutual information can produce coordinated behaviours for large robotic groups. The results obtained indicate that some of the evolved strategies scale well to significantly larger groups. In future work, we also plan to run evolutionary experiments with larger groups and with groups of varying size. In fact, variation in the group size might channel the evolutionary process toward the selection of scalable solutions.

6 Conclusions

In this paper, we investigated the use of information theoretic measures for the evolution of coordinated behaviours in groups of homogeneous robots. In particular, we defined a fitness function mainly based on the mean mutual information between the motor states of all possible robot pairs within a group of three robots. The results obtained show that evolution is able to find solutions that maximise the mutual information. This corresponds, in qualitative terms, to controllers that produce *structured* and *coordinated* behaviours. This is mainly the result of two different evolutionary drives. On the one hand, the maximisation of the mutual information corresponds to the maximisation of the individual entropy (see (6)). This favours the evolution of individual behaviours that allow the robot to visit different motor states during its lifetime. The embodiment of the robot, and the particular way we defined the computation of the motor state—as defined by (10) and (11)—favour the evolution of behaviours in which the motor state varies smoothly with time, producing sequences of atomic movements with varying time duration. These sequences are also periodic, due to the necessity of visiting as many motor states as possible for multiple times. On the other hand, the maximisation of the mutual information corresponds to the minimisation of the joint entropy between the motor states of two robots, which also corresponds to the observation of motor states that are positively correlated. The homogeneity of the robotic group ensures that this positive correlation results in synchronous behaviours.

[9] The situation is different for the behaviour evolved in run 13, which belongs to the same strategy. In this case, problems are given by the collision avoidance mechanism which exploits the difference between the self-emitted signal and the average group signal. With many robots, this difference is larger than with 3 robots only, a situation that was never experienced during evolution. As a consequence, not being able to avoid collisions, robots score a null performance.

We presented the results of two experiments that differ mainly in the characteristics of the environment, which may or may not offer obvious regularities to be exploited for coordination among the robots. We observed that, when these regularities are present, artificial evolution finds a way to exploit them to produce structured behaviours and to support the achievement of coordination among the robots. The situation is more complicated when such environmental regularities are removed. In this case, we showed how robots rely on the social environment, either by aggregating in one place and exploiting visual information or by relying on communication only. In the latter case, it is interesting to notice how synchronous behaviours can be performed without any reference to the position and orientation of other robots. We also presented the results of tests with physical robots, which should be considered a proof-of-concept of the applicability of the proposed methodology to real world scenarios. These tests demonstrate that several of the controllers evolved in simulation work also with physical robots (12 out of 18 in the E_l setup, 5 out of 11 in the E_d setup), and prove that the methodology we propose can produce controllers that reliably function in the real world. In other words, the results obtained are not based on unrealistic assumptions or on the exploitation of simulated features—such as uniform white noise—that may not have a correspondence in the real world. Instead, entropy and mutual information are maximised by the sensory–motor coordination of the robots, and by the synchronisation of the individual behaviours. Finally, we presented a scalability analysis of the evolved behaviours. This analysis proves the applicability of the proposed methodology also to large groups of robots, and paves the way towards efficient evolution of self-organising behaviours for robotic swarms.

We believe that the proposed methodology is particularly relevant for swarm robotics research, as it can efficiently synthesise self-organising, coordinated behaviours for a robotic swarm. In fact, there is a fundamental problem—referred to as the *design problem*—that arises in the development of self-organising behaviours for a group of robots (see also Funes et al. 2003; Trianni et al. 2008, for a detailed discussion of this topic). This problem consists in defining the appropriate individual rules that will lead to a certain global pattern, and it is particularly challenging due to the indirect relationship between control rules and individual behaviour, and between interacting individuals and the desired global pattern. In this respect, evolutionary robotics is particularly suited to synthesise self-organising behaviours. In fact, it bypasses the design problem as it relies on the automatic generation, test and selection of control solutions for the robotic system as a whole, without the need of an arbitrary decomposition of the given control problem into sub-problems (e.g., the desired global behaviour into individual behaviours and inter-individual interactions, as well as the individual behaviour in a set of control rules). The methodology we propose in this paper goes a step further in this direction: it promotes the evolution of coordinated behaviours without any constraint imposed by an explicit description of the desired solution. As a consequence, the proposed approach does not require a thorough knowledge of the system under study to devise the individual control rules, neither does it need a description of the desired solution to promote cooperative behaviours, as it can benefit of a task-independent, implicit utility function.

The proposed methodology represents a first step towards the evolution of self-organising behaviours for robotic swarms. In future work, we plan to exploit information theoretic measures in support of the evolution of task-oriented group behaviours. So far, we obtained synchrony without any constraint on the characteristics of the individual behaviour. We believe that a task-independent function can be successfully used in combination with a task-oriented one (on this issue, see also Prokopenko et al. 2006a). The former should provide the drives to synthesise structured and coordinated behaviour. The latter should simply channel the evolutionary process towards individual and group behaviours that serve specific functionalities. Another possible extension over the work presented in this paper concerns the use of heterogeneous robots. Using different controllers and/or different sensory–motor apparatus, it should be possible to observe coordination among the robots that does not forcedly limit to synchronisation of the movements. Turn taking, entrainment and other forms of coordination become possible whenever the robots may have access to different sensory–motor experiences. Finally, we intend to investigate also the possibility of exploiting different information theoretic measures and different neural controllers, such as recurrent neural networks.

Acknowledgements This work was supported by the Swarmanoid and by the ECAgents projects. Both projects are funded by the Future and Emerging Technologies programme (IST-FET) of the European Commission, under grants IST-022888 and 001940, respectively. The information provided is the sole responsibility of the authors and does not reflect the European Commission's opinion. The European Commission is not responsible for any use that might be made of data appearing in this publication.

References

Baldassarre, G., Trianni, V., Bonani, M., Mondada, F., Dorigo, M., & 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.

Brenner, N., Bialek, W., & de Ruyter van Steveninck, R. (2000). Adaptive rescaling maximizes information transmission. *Neuron*, *26*, 695–702.

Capdepuy, P., Polani, D., & Nehaniv, C. (2007). Maximization of potential information flow as a universal utility for collective behaviour. In *Proceedings of the 2007 IEEE symposium on artificial life (CI-ALife 2007)* (pp. 207–213). Piscataway: IEEE Press.

Cianci, C. M., Raemy, C., Pugh, J., & Martinoli, A. (2007). Communication in a swarm of miniature robots: the e-puck as an educational tool for swarm robotics. In E. Şahin, W. M. Spears, & A. F. T. Winfield (Eds.), *Lecture notes in computer science: Vol. 4433. Swarm robotics—second SAB 2006 international workshop, revised selected papers* (pp. 103–115), Rome, Italy, September 30–October 1, 2006. Berlin: Springer.

Feldman, D. (2002). *A brief introduction to: information theory, excess entropy and computational mechanics*. (Technical report). College of the Atlantic, Bar Harbor, ME.

Funes, P., Orme, B., & Bonabeau, E. (2003). Evolving emergent group behaviors for simple humans agents. In W. Banzhaf, T. Christaller, P. Dittrich, J. T. Kim, & J. Ziegler (Eds.), *Lecture notes in artificial intelligence: Vol. 2801. Advances in artificial life. Proceedings of the 7th European conference on artificial life (ECAL 2003)* (pp. 76–89). Berlin: Springer.

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

Klyubin, A., Polani, D., & Nehaniv, C. (2005a). All else being equal being empowered. In M. Capcarrere, A. A. Freitas, P. J. Bentley, C. G. Johnson, & J. Timmis (Eds.), *Lecture notes in artificial intelligence: Vol. 3630. Advances in artificial life. Proceedings of the 8th European conference on artificial life (ECAL 2005)* (pp. 744–753). Berlin: Springer.

Klyubin, A., Polani, D., & Nehaniv, C. (2005b). Empowerment: a universal agent-centric measure of control. In *Proceedings of the 2005 IEEE congress on evolutionary computation* (pp. 128–135). Piscataway: IEEE Press.

Lungarella, M., & Pfeifer, R. (2001). Robots as cognitive tools: Information theoretic analysis of sensory-motor data. In *Proceedings of the 2nd international IEEE/RSJ conference on humanoid robotics* (pp. 245–252). Piscataway: IEEE Press.

Lungarella, M., & Sporns, O. (2005). Information self-structuring: key principle for learning and development. In *Proceedings of the 4th international conference on development and learning* (pp. 25–30). Piscataway: IEEE Press.

Lungarella, M., Pegors, T., Bulwinkle, D., & Sporns, O. (2005). Methods for quantifying the information structure of sensory and motor data. *Neuroinformatics*, *3*(3), 243–262.

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

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

Mondada, F., & Bonani, M. (2007). *The e-puck education robot*. http://www.e-puck.org/.

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

Olsson, L., Nehaniv, C., & Polani, D. (2005). Sensor adaptation and development in robots by entropy maximization of sensory data. In *Proceedings of the 6th IEEE international symposium on computational intelligence in robotics and automation (CIRA-2005)* (pp. 587–592). Piscataway: IEEE Computer Society.

Prokopenko, M., & Wang, P. (2003). Evaluating team performance at the edge of chaos. In D. Polani, B. Browning, A. Bonarini, & K. Yoshida (Eds.), *Lecture notes in computer science: Vol. 3020*. *RoboCup 2003: robot soccer world cup VII* (pp. 89–101). Berlin: Springer.

Prokopenko, M., Gerasimov, V., & Tanev, I. (2006a). Evolving spatiotemporal coordination in a modular robotic system. In S. Nolfi, G. Baldassarre, R. Calabretta, J. C. T. Hallam, D. Marocco, J. A. Meyer, O. Miglino, & D. Parisi (Eds.), *From animals to animats 9: 9th international conference on the simulation of adaptive behavior (SAB 2006)* (pp. 558–569). Berlin: Springer.

Prokopenko, M., Gerasimov, V., & Tanev, I. (2006b). Measuring spatiotemporal coordination in a modular robotic system. In L. M. Rocha, L. S. Yaeger, M. A. Bedau, D. Floreano, R. L. Goldstone, & A. Vespignani (Eds.), *Artificial life X: proceedings of the tenth international conference on the simulation and synthesis of living systems* (pp. 185–191). Cambridge: MIT Press.

Quinn, M., Smith, L., Mayley, G., & 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.

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

Sporns, O., & Lungarella, M. (2006). Evolving coordinated behavior by maximizing information structure. In L. M. Rocha, L. S. Yaeger, M. A. Bedau, D. Floreano, R. L. Goldstone, & A. Vespignani (Eds.), *Artificial life X: proceedings of the tenth international conference on the simulation and synthesis of living systems* (pp. 323–329). Cambridge: MIT Press.

Sporns, O., Tononi, G., & Edelman, G. (2000). Connectivity and complexity, the relationship between neuroanatomy and brain dynamics. *Neural Networks*, *13*, 909–922.

Tarapore, D., Lungarella, M., & Gomez, G. (2004). Fingerprinting agent-environment interaction via information theory. In F. Groen, N. Amato, A. Bonarini, E. Yoshida, & B. Kröse (Eds.), *Intelligent autonomous systems 8* (pp. 512–520). Amsterdam: IOS.

Tarapore, D., Lungarella, M., & Gomez, G. (2006). Quantifying patterns of agent-environment interaction. *Robotics and Autonomous Systems*, *54*(2), 150–158.

Tononi, G., Sporns, O., & Edelman, G. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences*, *91*, 5033–5037.

Tononi, G., Sporns, O., & Edelman, G. G. (1996). A complexity measure for selective matching of signals by the brain. *Proceedings of the National Academy of Sciences*, *93*, 3422–3427.

Tononi, G., Edelman, G., & Sporns, O. (1998). Complexity and coherency: integrating information in the brain. *Trends in Cognitive Sciences*, *2*(12), 474–484.

Trianni, V., & Nolfi, S. (2007). Minimal communication strategies for self-organising synchronisation behaviours. In *Proceedings of the 2007 IEEE symposium on artificial life (CI-ALife 2007)* (pp. 199–206). Piscataway: IEEE Press.

Trianni, V., Nolfi, S., & Dorigo, M. (2008). Evolution, self-organisation and swarm robotics. In C. Blum & D. Merkle (Eds.), *Natural computing series*. *Swarm intelligence. Introduction and applications*. Berlin: Springer.

Van Dyke Parunak, H., & Brueckner, S. (2001). Entropy and self-organization in multi-agent systems. In *Proceedings of the fifth international conference on autonomous agents* (pp. 124–130). New York: ACM.