1

Mutual Information as a task-independent utility function for evolutionary robotics.

Valerio Sperati, Vito Trianni, and Stefano Nolfi

Laboratory of Autonomous Robotics and Artificial Life
Institute of Cognitive Sciences and Technologies, CNR
via S. Martino della Battaglia, 44 - 00185 Rome, Italy

1.1 Introduction

The design of the control system for a swarm of robots is not a trivial enterprise. Above all, it is difficult to define which are 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 task-independent 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 the 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.

In this paper, we investigate whether information-theoretic measures can be used to drive the evolution of coordinated behaviours in groups of evolving robots. In particular, we demonstrate how the use of a utility function that maximises the mutual information 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. Looking at the individual level, 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. Looking at the collective level, we define a "coordinated" behaviour as a situation in which the behaviours produced by the individuals are correlated as for example, in the case of the production of similar and synchronized oscillatory movements or as in the case of alternated turn-taking behaviours.

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 to provide exploitable environmental cues, 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—non trivial—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 Sec. 1.3 we briefly review related literature. In Sec. 1.4 and 1.5, we describe the experimental setup and the results obtained. Finally, in Sec. 1.6 we discuss the main contributions of the paper and we draw our conclusions.

1.2 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 Sec. 1.4.3. Regarding notation, we follow Feldman's style: we use capital letters to indicate a random variable, and lowercase letters to indicate a particular value of that 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:

$$H[X] = -\sum_{x \in \mathcal{X}} p(x) \cdot \log_2\ p(x). \tag{1.1}$$

The entropy $H[X]$ - or *marginal entropy* - 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 in *alphabet* $\mathcal{X}$ with the same probability ($\frac{1}{M}$). There are many interpretations about 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:

$$H[XY] = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(xy)\cdot\log_2\ p(xy). \tag{1.2}$$

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:

$$H[X|Y] = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(xy)\cdot\log_2 p(x|y). \tag{1.3}$$

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 re-expressed in terms of marginal entropy and conditional entropy:

$$H[XY] = H[X] + H[Y|X] = H[Y] + H[X|Y]. \tag{1.4}$$

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

$$MI[X;Y] = -\sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(xy)\cdot\log_2\frac{p(x)\cdot p(y)}{p(xy)}. \tag{1.5}$$

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

$$MI[X;Y] = H[X] + H[Y] - H[XY]. \tag{1.6}$$
$$MI[X;Y] = H[X] - H[X|Y] = H[Y] - H[Y|X]. \tag{1.7}$$

Looking at eq. (1.5), it is possible to notice that $MI[X;Y] = 0$ if the two variables are statistically independent. On the other hand, eq. (1.6) shows that $MI[X;Y] = 0$ if the two variables have zero entropy. The measure is symmetric, namely $MI[X;Y] = MI[Y;X]$.

The interpretation of MI is quite clear looking at eq. (1.7). 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. When this happens, we can imagine that a bidirectional communication channel through whom the information flows, establishes between the two variables. In practice, MI can be used as a powerful index of correlation: the greater the value of MI, the more correlate are two variables. The great advantage of MI is that it takes into account both linear and nonlinear dependencies (Lungarella and Pfeifer 2001).

1.3 Related Work

The above measures and related derivations have been successfully used as analytic tools in different fields. In ethology for example, information theory was used to describe the interplay between pheromone molecules and ants' movements. By observing ants' foraging behaviour, Van Dyke Parunak and Brueckner (2001) showed that the increase in entropy at the micro-level of the chemical particles is compatible with the 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 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 robot can be treated in an efficient and computationally economic way if sensors can adapt to the statistical properties of the incoming signals. Tarapore et al. (2004, 2006) applied entropy and mutual information to the sensory channels of a two wheeled simulated robot: These measures are 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 (e.g., foveation on a red object), with uncoordinated ones (e.g., 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 work. The authors argued that coordinated sensory-motor activity induces information structures in the sensory experience.

This idea has been further elaborated by Klyubin et al. (2005a) with the notion of *Empowerment*, an information based quantity that allows to characterise the efficiency of the perception-action loop of an organism model. This quantity measures the potential of the organism to imprint information on the environment via its actuators in a way that can be recaptured by its sensors. A generalization of this measure for continuous domains has been later proposed by Jung et al. (2011).

Finally, other recent works investigated the use of information-based measures to characterise collective behaviour. Wang et al. (2011) investigate how information propagate in groups of coordinated individuals. In particular the authors showed how crucial phases in collective behaviours (corresponding to clustering, merging, and separation) are characterised by well-marked peaks of the active information storage and transfer entropy measures. Harder et al. (2010) instead demonstrated how the mutual information among the actions displayed by a couple of coordinated individuals can be used to characterise the autonomy level of the individuals and whether the couple can be characterised as a unique coherent entity. Harder et al. (2011) demonstrated how information theoretic measures can be used to quantify the ability of individual agents to extract information locally about global features. Finally, information theoretic measures have been used to characterise the emergence of self-organizing collective properties, e.g. the abrupt formation of a dynamic chain pattern within a swarm of robots evolved for the ability to navigate between two target areas (Sperati et al. 2011).

More importantly, from the point of view of the objectives of this paper, several recent works demonstrated how information theoretic measures can be used to synthesize sensory-motor coordination capabilities and/or co-ordinated behaviours. 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 selecting 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. Prokopenko et al. (2006) demonstrated how the maximisation of *Excess Entropy* (a measure of apparent memory or structure in a system), can lead to useful coordinated behaviour. In particular, the authors showed how a simulated snake-like modular robot, evolved on the basis of this measure, displays an effective locomotion behaviour: the linked actuators composing the robot get coordinated and produce a forward motion which interestingly adapts to the environment features, and makes the robot capable to face challenging terrains characterised by obstacles, narrow passages and ragged textures. In a subsequent work involving the same robotic setup, the authors used as fitness function the *Transfer Entropy*, i.e. a measure of information transfer (Lizier et al. 2008). In the experiment reported in this paper the authors observed a propagation of information between the head and the tail of the robot. The observed information transfer structures are analogous to gliders in cellular automata, which have been demonstrated to represent the coherent transfer of information across space and time, and play an important role in facilitating distributed computation.

Zahedi et al. (2010) proposed a learning method based on the maximisation of the predictive information, i.e. the mutual information between past and future sensors states, in the sensory-motor loop. The method was evaluated on a series of experiments involving robot constituted by chains of individually controlled elements of varying length. The *Empowerment* measure mentioned above has been used as a selection criteria for evolving the properties of the agent sensors and actuators (Klyubin et al. 2008, 2005b). As observed by the authors this triggers a process in which the morphology of the sensors and of the actuators adapts to the characteristics of the environment in which the agents are situated. More generally the authors claim that empowerment, being a measure of what the agents could do rather than a measure of what they actually do, can constitute a general adaptive drive that could enable the development of survival-relevant behaviour even in the absence of behaviour specific drives (Klyubin et al. 2008). Empowerment and *Infotaxis*, a measure that encodes the expected reduction of entropy achieved by selecting actions, have also been successfully used to synthesize coordinated collective behaviours (Capdepuy et al. 2007; Salge and Polani 2011).

Most of these works—as much as the study presented in this chapter—belong to a novel methodology in evolutionary robotics called *information-driven evolution*, in which information based measures that are task-independent are used as utility functions.

1.4 Experimental setup

As mentioned in Sec. 1.2 , $MI[X;Y]$ can be seen as a powerful measure to grasp the correlation between two stochastic processes X and Y. Moreover, maximising MI also corresponds to maximising the entropy of the single processes $H[X]$ 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 (see also Sperati et al. (2008)). 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 processes 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 Section 1.4.3). In particular, we focus on the wheels' speed, which characterise 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 Section 1.4.1). Robots are placed in a square arena of 1x1 m in side 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. 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.

[1] This is true if the joint entropy is kept constant, see eq. (1.6).

Fig. 1.1. Left: The e-puck robot developed at the EPFL in Lausanne, Switzerland (Mondada and Bonani 2007). Right: A close up view of the environment with the light bulb in the centre and three robots.

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 lasts 200 generations. In each generation, the population is evaluated and genotypes are selected for reproduction on the basis of an estimate of their fitness (see Section 1.4.2). This estimate is obtained by testing each genotype 10 times—i.e., we perform 10 independent *trials* randomly varying the initial conditions (see Section 1.4.3). The best evolved genotypes resulting from each evolutionary run is then selected for a qualitative and quantitative analysis, presented in Section 1.5.

1.4.1 The Robot and the Neural Controller

The experiments presented in this paper are performed using the *e-puck* robots (see Fig. 1.1 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 in the following sections. 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 which are 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 leave the evolutionary process free to explore a wide set of possible solutions.

Each robot is provided with various sensory systems 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 the camera of the other robots. Additionally, robots can exploit their wireless bluetooth interface to send and receive short messages (see Fig. 1.2).

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

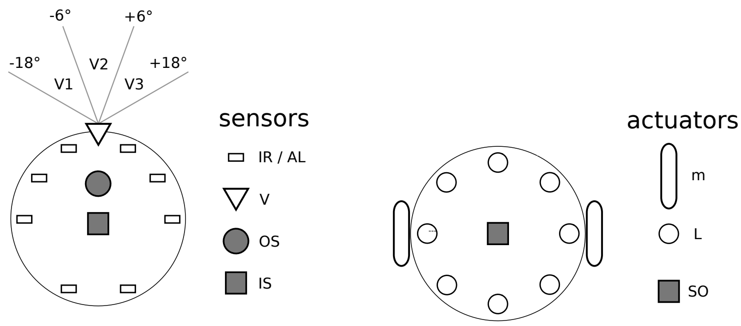

Fig. 1.2. Robot sensors and actuators. Left: sensors “IR” (8 infrared sensors), “AL” (8 light sensors), ”V” (3 fields of view from pre-processed camera data), “IS” (average group signal), “OS” (own signal). Right: actuators “m” (wheel velocity and direction), ”L” (leds ring on/off), “SO” (signal). Note that gray symbols refer to virtual actuators/sensors not present in the physical robot, but implemented through the wireless bluethoot interface.

with an identical control structure—a homogeneous group. Each robot is controlled by a fully connected two layer neural network with fixed topology, characterised by an input layer with leaky integrator neurons and by an output layer of motor neurons (see Fig. 1.3). The activation of the output neurons is computed as the weighted sum of all input units and the bias, filtered through a sigmoid function:

$$O_j(t) = \sigma\left(\sum_i w_{ij} I_i(t) + \beta_j\right), \qquad \sigma(z) = \frac{1}{1+e^{-z}}, \tag{1.8}$$

where $I_i(t)$ corresponds to the activation of the i^{th} 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:

$$I_i(t) = \tau_i \cdot I_i(t-1) + (1-\tau_i) \cdot \mathcal{I}_i(t), \tag{1.9}$$

where τ_i is the time constant of the i^{th} 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. 1.3). 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 ($IF_i, i = 1, \ldots, 8$), which can detect an obstacle up to a distance of approximately 25 mm (see Fig. 1.3). Three sensory inputs ($V_i, i = 1, \ldots, 3$) 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 ($AL_i, i = 1, \ldots, 8$), shown in grey in Fig. 1.3.

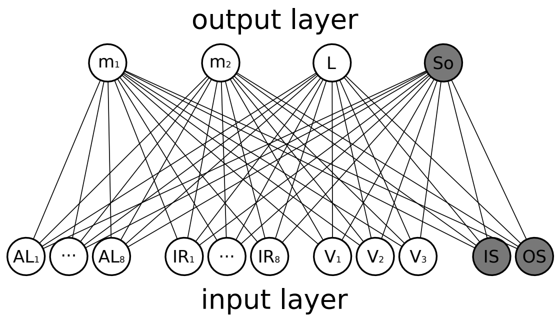

Fig. 1.3. The architecture of the neural controller. Note that gray neurons refers to virtual actuator/sensors not present in the physical robot, but implemented through the wireless bluethoot interface. Neurons "AL" are used in the experiment E_l only, while the other neurons are common to both setups.

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).

1.4.2 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 three identical neural controllers which are downloaded onto three identical robots (i.e., the robots are homogeneous). Each team 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 eq. (1.12), 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.

1.4.3 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.

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:

$$M_j(t) = \gamma \cdot M_j(t-1) + (1-\gamma) \cdot m_j(t), \qquad j = 1, 2 \tag{1.10}$$

[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.

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:

$$X = \lfloor M_1 \cdot 5 \rfloor + \lfloor M_2 \cdot 5 \rfloor \cdot 5, \tag{1.11}$$

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 (1.5). 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:

$$F = \frac{\sum_{i=1}^{N} \sum_{j=i+1}^{N} MI[X_i; X_j] \cdot \frac{20-c}{20}}{a} \qquad a = \binom{N}{2} \tag{1.12}$$

where N is the number of robots, c is the number of times in which one of the robots collided against a wall or against another robot, truncated to 20, a is the binomial coefficient for couples of robots. In other worlds, this equation computes the average Mutual Information calculated between all possible pairs of 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.

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 practice. The main reason for this is that the individual entropy cannot be maximised because robots are embodied and their dynamical interaction

[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 us 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.

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.

1.5 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.

1.5.1 Experiment E_l

In the experiment E_l, the robots are situated in in square arena of 1x1 m in side 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 eq. (1.12). The results obtained are summarised in Table 1.1, in which we show mean and standard deviation over the 500 trials of the fitness F, of the average mutual information $\widehat{MI}$ over all possible robot pairs, and of the average 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, respectively obtained in run 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 1.1). 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.

Behavioural Analysis

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

Table 1.1. Experiment E_l: fitness F, average mutual information $\widehat{MI}$ and average 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

by the 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 1.1). 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. 1.4 and the video "replication 16 E_l" in the online supplementary material [5]). 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. 1.4 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, the left motor activations also 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.

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. 1.4 right). In short, robots are synchronised as they perform the same movements at the same time. The mechanism

[5] See `http://laral.istc.cnr.it/esm/sperati-etal-GSO_2012.html` for videos and other supplementary material.

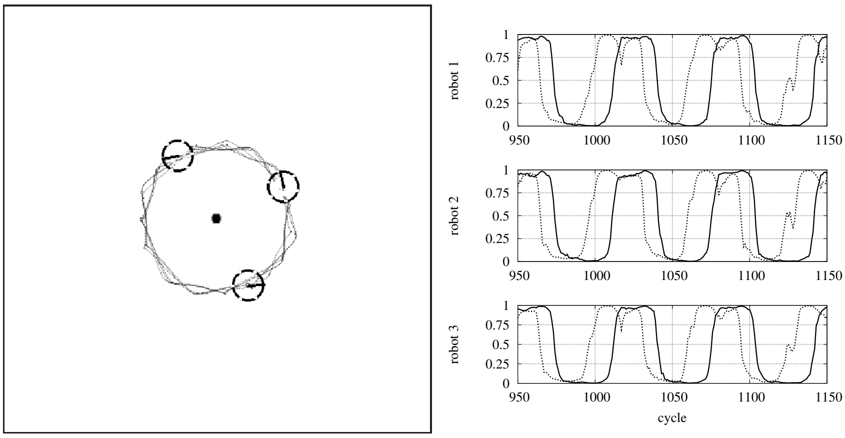

Fig. 1.4. Analysis of the behaviour produced by the best evolved controller in run 16 of experiment E_l. Left: trajectories of the robots. Right: activation of the motor neurons of each robot, plotted from cycle 950 to cycle 1150 to highlight the periodic motion of the robots. The solid and dotted lines indicate respectively the left and right motor neurons.

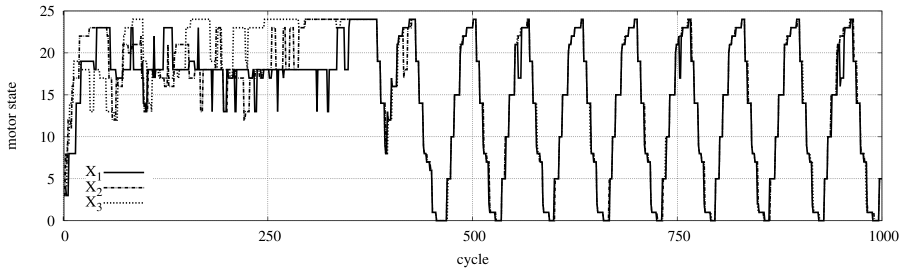

Fig. 1.5. The motor states of the three robots—computed using eq. (1.11)—are plotted against the number of cycles. Notice the initial coordination phase, followed by synchronised movements.

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 (2009) —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 (2009) —waiting for the other robots. Synchronised movements start when all robots achieve the reset configuration.

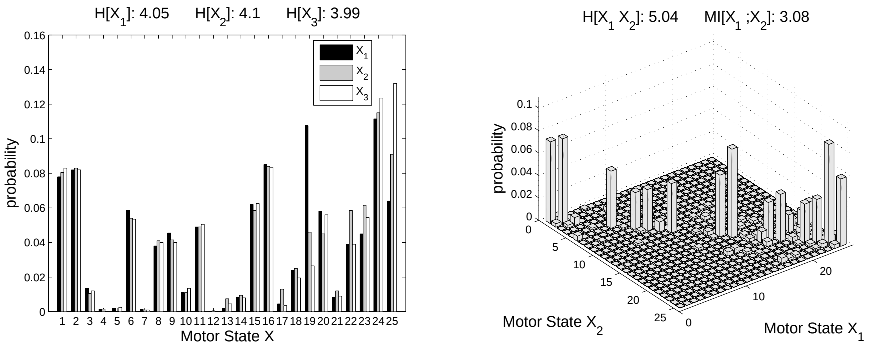

Fig. 1.6. Left: Probability distribution for the motor states X_i of each robot $i = 1, 2, 3$. Right: Probability distribution of the joint state $\langle X_1, X_2 \rangle$.

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 1.5 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 take on many different values. In other words, the motor states of the robots vary considerably through 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. 1.6. 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. 1.6, 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 $H[X_1X_2]$, 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.

[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.

Table 1.2. Experiment E_l: average mutual information ($\widehat{MI}$) and average 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 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

run	$\widehat{MI}$	$\widehat{H}$	ratio
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

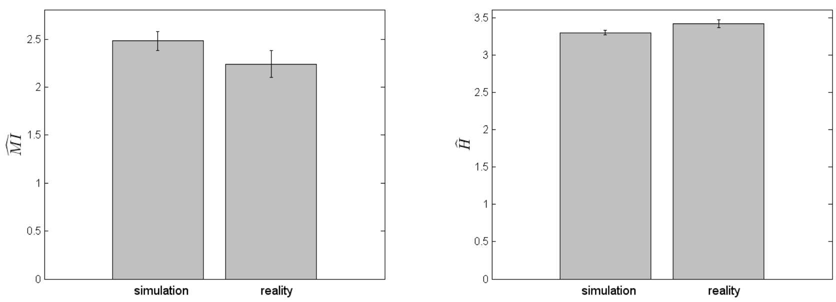

Fig. 1.7. Average mutual information ($\widehat{MI}$) and entropy ($\widehat{H}$) computed by testing the best evolved controller of run 9 of experiment E_l in simulation and in reality for 20 trials of 2000 cycles. During the tests in hardware, the robots were situated in the same randomly generated positions and orientations that were used for the tests in simulation.

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.

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 average mutual information computed among all possible robot pairs. The results obtained are shown in Table 1.2, along with the ratio with the average mutual information resulting from simulation. It is worth noting that the ratio between the mutual information

[7] See videos in the online supplementary material

observed in simulation and in the real environment is generally quite high, indicating that the behaviours tested in reality correspond fairly well to those observed in simulation.

After this preliminary test was 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 [8]). 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. 1.7 left. Similarly, the mean entropy over 20 trials computed on the tests with physical robots corresponds to the value obtained in simulation (see Fig. 1.7 right).

1.5.2 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 1.3, 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 respectively in run 14 and 10. 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 non-structured 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.

[8] See video “replication 9 E_l” in the online supplementary material.

Table 1.3. Experiment E_d: fitness F, average mutual information $\widehat{MI}$ and average 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

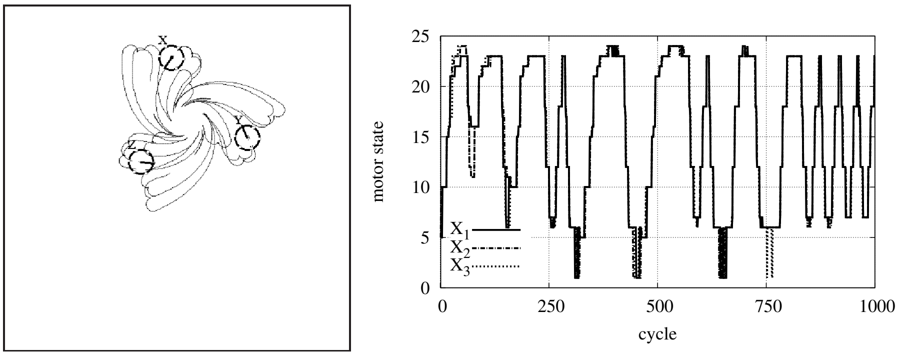

Fig. 1.8. Left: Trajectories of the robots produced by the best evolved controller in run 6 of experiment E_d. Right: The motor states of the three robots are plotted against the number of cycles.

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 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.

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. 1.8 left and video "replication 6 E_d" in the online supplementary

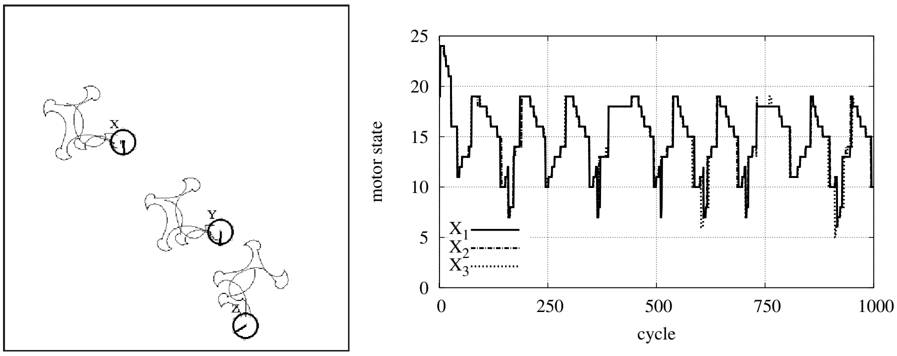

Fig. 1.9. Left: Trajectories of the robots produced by the best evolved controller in run 10 of experiment E_d. Right: The motor states of the three robots are plotted against the number of cycles.

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 to one 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.1.8 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 trial, as can be noticed in Figure 1.8 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. 1.9 and video "replication 10 E_d" 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, 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 from the very beginning of the trial, because there is no need to achieve a particular spatial formation (see the motor states plotted in Fig.1.9 right).

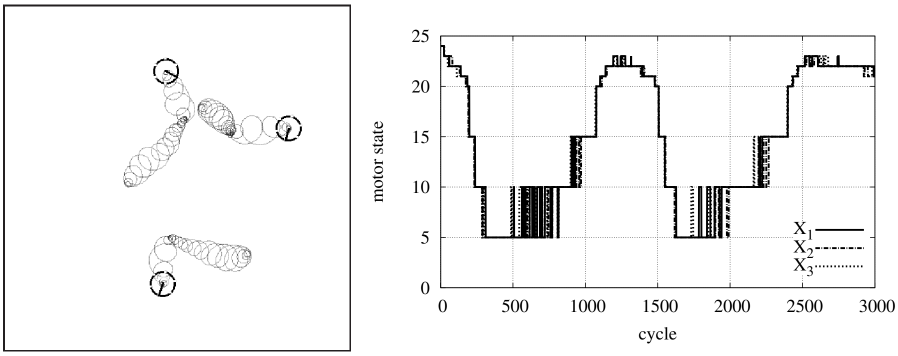

Fig. 1.10. Left: Trajectories of the robots produced by the best evolved controller in run 8 of experiment E_d. Right: The motor states of the three robots are plotted against the number of cycles.

Finally, the last strategy includes only the controller evolved in run 8 (see Fig. 1.10 and video "replication 8 E_d" 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.1.10 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 this behaviour (see the coordinated motor states in Fig.1.10 right), and it is achieved and maintained exploiting communication only.

Porting to Reality

By testing with physical robots all controllers that proved successful in simulation, we observed good generalisation only in 5 out of 11 cases, namely runs 2, 8, 9, 10, 13[9] 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.

[9] see videos in the online supplementary material.

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 1.4).

Table 1.4. Experiment E_d: average mutual information ($\widehat{MI}$) and average 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	$2,69 \pm 0,06$	$3,36 \pm 0,07$	0.99
8	$1,91 \pm 0,03$	$2,65 \pm 0,06$	0.94
9	$2,06 \pm 0,10$	$3,21 \pm 0,26$	0.95
10	$2,88 \pm 0,05$	$3,51 \pm 0,06$	0.98
13	$1,66 \pm 0,06$	$2,90 \pm 0,20$	0.92

1.6 Conclusion

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 based on the average 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 eq. (1.6)). This favours the evolution of individual behaviours that allow the robot to produce different actions during its lifetime. The embodiment of the robot, and the particular way we defined the computation of the motor state—as defined by eq. (1.10) and (1.11)—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 to visit 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 leads to coordinated synchronous behaviours.

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 the environment does not provide such regularities. In this case, the robots exploit the possibility to generate the required regularities through social behaviours (i.e. by aggregating and/or by communicating). Moreover we observed how the obtained results can be validated in hardware. More specifically, we demonstrated how 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). Overall this demonstrates how the proposed measure is able to synthesize robust solution that can overcome the problems caused the simulation-reality gap.

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 (Trianni and Nolfi 2012). 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. 2006). 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.

1.7 Acknowledgments

This research has been supported by CNR under the European Science Foundation project Hierarchical Heterogeneous SWARM (H2Swarm).

References

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

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

Capdepuy, P., Polani, D., and 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)*, pages 207–213. IEEE Press, Piscataway, NJ.

Cianci, C. M., Raemy, C., Pugh, J., and Martinoli, A. (2007). Communication in a swarm of miniature robots: The e-puck as an educational tool for swarm robotics. In Şahin, E., Spears, W. M., and Winfield, A. F. T., editors, *Swarm Robotics - Second SAB 2006 International Workshop, Rome, Italy, September 30-October 1, 2006 Revised Selected Papers*, volume 4433 of *Lecture Notes in Computer Science*, pages 103–115. Springer Verlag, Berlin, Germany.

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., and Bonabeau, E. (2003). Evolving emergent group behaviors for simple humans agents. In Banzhaf, W., Christaller, T., Dittrich, P., Kim, J. T., and Ziegler, J., editors, *Advances in Artificial Life. Proceedings of the 7th European Conference on Artificial Life (ECAL 2003)*, volume 2801 of *Lecture Notes in Artificial Intelligence*, pages 76–89. Springer Verlag, Berlin, Germany.

Harder, M., Polani, D., and Nehaniv, C. (2010). Two agents acting as one. In *Proc. of the Alife XII Conference*.

Harder, M., Polani, D., and Nehaniv, C. (2011). Think globally, sense locally: From local information to global features. In *IEEE Symposium on Artificial Life*.

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

Jung, T., Polani, D., and Stone, P. (2011). Empowerment for continuous agent-environment systems. *Adaptive Behavior*, 19:16–39.

Klyubin, A., Polani, D., and Nehaniv, C. (2005a). All else being equal being empowered. In Capcarrere, M., Freitas, A. A., Bentley, P. J., Johnson, C. G., and Timmis, J., editors, *Advances in Artificial Life. Proceedings of the 8th European Conference on Artificial Life (ECAL 2005)*, volume 3630 of *Lecture Notes in Artificial Intelligence*, pages 744–753. Springer Verlag, Berlin, Germany.

Klyubin, A., Polani, D., and Nehaniv, C. (2005b). Empowerment: A universal agent-centric measure of control. In *Proceedings of the 2005 IEEE Congress on Evolutionary Computation*, pages 128–135. IEEE Press, Piscataway, NJ.

Klyubin, A., Polani, D., and Nehaniv, C. (2008). Keep your options open: An information-based driving principle for sensorimotor systems. *PLoS ONE*, 3(12).

Lizier, J., Prokopenko, M., Tanev, I., and Zomaya, A. (2008). Emergence of glider-like structures in a modular robotic system. In *Proc. Eleventh International Conference on the Simulation and Synthesis of Living Systems (ALife XI)*, pages 366–373.

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

Lungarella, M. and 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*, pages 245–252. IEEE Press, Piscataway, NJ.

Lungarella, M. and Sporns, O. (2005). Information self-structuring: Key principle for learning and development. In *Proceedings of the 4th International Conference on Development and Learning*, pages 25–30. IEEE Press, Piscataway, NJ.

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

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

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

Nolfi, S. and Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MIT Press/Bradford Books, Cambridge, MA.

Olsson, L., Nehaniv, C., and 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)*, pages 587–592. IEEE Computer Society Press, Piscataway, NJ.

Prokopenko, M., Gerasimov, V., and Tanev, I. (2006). Evolving spatiotemporal coordination in a modular robotic system. In Nolfi, S., Baldassarre, G., Calabretta, R., Hallam, J., Marocco, D., Meyer, J., Miglino, O., and Parisi, D., editors, *From Animals to Animats 9: 9th International Conference on the Simulation of Adaptive Behavior (SAB 2006)*, pages 558–569. Springer Verlag, Berlin, Germany.

Prokopenko, M. and Wang, P. (2003). Evaluating team performance at the edge of chaos. In Polani, D., Browning, B., Bonarini, A., and Yoshida, K., editors, *RoboCup 2003: Robot Soccer World Cup VII*, volume 3020 of *Lecture Notes in Computer Science*, pages 89–101. Springer Verlag, Berlin, Germany.

Quinn, M., Smith, L., Mayley, G., and Husbands, P. (2003). Evolving controllers for a homogeneous system of physical robots: Structured cooperation with minimal sensors. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, 361:2321–2344.

Salge, C. and Polani, D. (2011). Local information maximisation creates emergent flocking behaviour. In *Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on Artificial Life*.

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

Sperati, V., Trianni, V., and Nolfi, S. (2008). Evolving coordinated group behaviours through maximisation of mean mutual information. *Swarm Intelligence*, 2(2–4):73–95.

Sperati, V., Trianni, V., and Nolfi, S. (2011). Self-organised path formation in a swarm of robots. *Swarm Intelligence*, 5(2):97–119.

Sporns, O. and Lungarella, M. (2006). Evolving coordinated behavior by maximizing information structure. In Rocha, L., Yaeger, L., Bedau, M., Floreano, D., Goldstone, R., and Vespignani, A., editors, *Artificial Life X: Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems*, pages 323–329. MIT Press, Cambridge, MA.

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

Tarapore, D., Lungarella, M., and Gomez, G. (2004). Fingerprinting agent-environment interaction via information theory. In Groen, F., Amato, N., Bonarini, A., Yoshida, E., and Kröse, B., editors, *Intelligent Autonomous Systems 8*, pages 512–520. IOS Press, Amsterdam, The Netherlands.

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

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

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

Trianni, V. and Nolfi, S. (2009). Self-organising sync in a robotic swarm. a dynamical system view. *IEEE Transactions on Evolutionary Computation*, 13(4):722–741.

Trianni, V. and Nolfi, S. (2012). *The Handbook of Collective Robotics - Fundamentals and Challenges*, chapter Evolving collective control, cooperation and distributed cognition, pages 168–189. Pan Stanford Publishing, Singapore.

Trianni, V., Nolfi, S., and Dorigo, M. (2008). Evolution, self-organisation and swarm robotics. In Blum, C. and Merkle, D., editors, *Swarm Intelligence. Introduction and Applications*, Natural Computing Series. Springer Verlag, Berlin, Germany.

Van Dyke Parunak, H. and Brueckner, S. (2001). Entropy and self-organization in multi-agent systems. In *Proceedings of the Fifth International Conference on Autonomous Agents*, pages 124–130. ACM Press, New York, NY.

Wang, X., Miller, J., Lizier, J., Prokopenko, M., and Rossi, L. (2011). Measuring information storage and transfer in swarms. In *Proc. Eleventh European Conference on the Synthesis and Simulation of Living Systems (ECAL 2011)*, pages 838–845.

Zahedi, K., Ay, N., and Der, R. (2010). Higher coordination with less control - a result of information maximization in the sensorimotor loop. *Adaptive Behavior*, 18:338–355.