International Journal of Combinatorial Optimization Problems and Informatics, 12(1), Jan-April 2021, 29-41. ISSN: 2007-1558.

www.editada.org

Evolution of Basic Communication Strategies in Artificial Agents

Víctor Hugo Santos[1], Manuel Pardo[1], Alejandra Ciria[2], Bruno Lara[1]

[1] Centro de Investigación en Ciencias, Universidad Autónoma del Estado de Morelos, México
[2] Facultad de Psicología, Univ. Nacional Autónoma de México, México
victor.santos@uaem.mx, jose.pardo@uaem.mx , ciriacontacto@gmail.com, bruno.lara@uaem.mx

Abstract. Communication and signaling processes are ubiquitous in biological organisms, having evolved throughout the evolutionary tree whenever transmission and receiving mechanisms are present. The work presented in this article addresses, from an evolutionary robotics approach, how a basic communication strategy can emerge in artificial agents through the co-evolution of signalers and perceivers. Although there are many works that model the evolution of the communication processes using autonomous robots, the work presented here attempts to tackle the problem using a different approach. Here, the emphasis is put on the perspective of the perceiver, by examining how the meaning of a signal emerges during the evolutionary process starting from a meaningless cue. This leads to effectively changing the behaviour of the perceiver whenever the signal is present. All this is achieved without directly codifying any behaviour related to perception in the fitness function of the evolutionary algorithm.

Keywords: Evolutionary robotics, communication, signaling, artificial agents.

Article Info
Received Sep 2, 2020
Accepted Dec 13, 2020

1 Introduction

Basic communication, signaling, is understood as the process, present in biological agents, that allows an agent (the perceiver) to make a decision or to take action, based on changes in the environment (signals) made by another agent (the emitter). These signals can be desired, when it is beneficial to the perceiver, undesired when it is detrimental or harmful to the perceiver, and, intended when is it beneficial to the emitter or unintended when is detrimental to the emitter [1, 2] .

The study of animal signaling is a biologically grounded discipline where the evolutionary theory underlies each and every communicative phenomenon [3]. The definition of animal signaling has been a matter of great debate in evolutionary biology (see [4]). The starting point of this discussion is often attributed to the seminal work by Dawkins and Kerbs [5] where the information-based definitions of communication were eliminated. This alternative perspective of animal signaling is an influence-based conception of morphological traits and behavioural actions with an evolved function to influence the behaviour of other individuals ([6, 7, 3]. The function of a signal is to influence behaviour, rather than the transmission of encoded information. Therefore, under this view, the terms ‘sender’ and ‘receiver’ have been replaced by ‘signaler’ and ‘perceiver’ to highlight a definition based on ‘signalers’ making an influence on ‘perceivers’ [3]. Within the proponents of an influence-based definition of animal signaling, there is no consensus if it requires either a benefit for the ‘signaler’ or a mutual benefit [8]. For the former, animal signaling definition should be signaler-centered because they receive the benefit of influencing perceivers and, thus, the perceiver benefit should be excluded. Thus, only the ‘signalers’ need a specialization through natural selection, shaping their behaviour and signals to influence perceivers via their evolutionary history of perceptual susceptibility [3]. For the later, animal signaling is understood as a co-adaptation, where signalers receive the benefit of influencing perceivers' behavior and perceivers are rewarded when responding to the signals [8, 7]. A signal influences the behaviour of other individuals because it evolved to have that effect, and it is effective because the perceiver’s response evolved to be affected by the signal. Adaptation thus needs to be used to define a signal, assuming that natural selection is sensitive to long-term benefits of a relationship [7].

Carazo [9] argued that information could not be eliminated from biological communication when the function and the information content of a signal are differentiated. They suggested a functional notion of information that is compatible with the adaptationism accounts of animal signaling. Functional information is what makes the influence of a signal adaptive for perceivers, being in a direct relation to the perceivers' fitness. Signalers do not aim to provide information, nonetheless, natural selection acted on perceivers’ ability to extract information from the signal and the context in which is [10]. Thus, signalers do not provide information, they trigger an effect that makes signaling adaptive for the signaler if the perceivers extract information to make adaptive decisions. Accordingly, the fitness of signalers is in a direct relation to the effect or effects of a signal on the perceiver. Communication may in a way be seen as exchanging information for effect, as perceivers need information if they are to display adaptive responses.

From a biological perspective, this type of communication has evolved millions of times, and it is present in any modality in which a biological organism can emit and perceive signals such as vision, hearing, electroreception, olfaction, touch, taste [1] or vibrations [2]. Furthermore, communication is present in all types of organisms [11] . Given the widespread of this phenomenon, there has to be a basic condition on the structure of the evolutionary process that can be modeled without being affected by the particularities of the agent, giving the results of the study a certain biological plausibility [12].

Agents should include artificial agents, converting communication into an ideal cognitive process to be modelled and studied from an evolutionary robotics perspective. Evolutionary robotics emerged in the 1990’s as a method to replicate, in artificial agents, the processes present in Darwinian evolution. Research in this area is based on the premise that the problems a situated autonomous robot (real or simulated) has to solve have no significant differences with a biological counterpart [13, 14]. Furthermore, evolutionary robotics can be used to test and model, from an evolutionary perspective, cognitive processes that by their characteristics cannot be studied in biological organisms. This can either be because of temporal constraints or because of lack of evidence.

Temporal constraints refer to the fact that the experimental study of evolution in biological organisms is very impractical in terms of the time it takes to replicate processes successfully. This applies even for species with a concise cycle of life, such as the iconic evolutionary experiments made by Lenski [15] with an e. coli culture. This on-going experiment, started in 1988, reported changes in the cultures after 2000 generations in 1991, after 3 years of research.

On the other hand, lack of evidence refers to the fact that for most cognitive processes there is no evolutionary physical evidence in the fossil record “[a]s cognitive adaptations do not fossilize” [16].

1.1. Evolutionary robotics

In evolutionary robotics (ER), The principles of Darwinian evolution are used as heuristic optimization processes. These principles have been used for two main purposes, namely, the evolution of bodies and the evolution of behaviours. Seminal work on the evolution of bodies and behaviours was first reported in [17] where they evolved basic robotic lifeforms capable of basic locomotion. In [18] the authors investigated the effects of development on evolutionary processes, reporting that morphological development increases evolvability in artificial lifeforms evolved for locomotion. The authors brought the process a step further developing a pipeline for the evolution of organisms in *silico*, which then were realized *in vivo* in [19]. The results of the evolutionary process were built using frog cells from different body locations.

In an attempt to evolve evolutionary adaptations that do not fossilize, very interesting results are found in [20], where the authors evolved the necessary controllers for agents to respond to basic emotions. In a very interesting work [21], the authors successfully evolved agents in different environments where prediction, by means of internal models, is a prerequisite for success in a tracking task. In [22] the authors use ER to look for the evolution of cooperation between individuals, finding that even when these are genetically related, this behavior is very difficult to evolve. In [23] the aim was to study the emergence of symbolic communication in artificial agents. Agents evolved basic communication schemes when these need to find a mate for reproduction. They suggest that this ability might have evolved without requiring previous high cognitive abilities, In a very interesting experiment, Kadish [24] explore the emergence of differentiation of sound emitted by different species of artificial agents. The authors use two different species of emitters and perceivers, which evolve to differentiate and successfully decode messages from their own species to those of others.

The research presented here uses the principles of ER to study the emergence of basic communication skills among simulated agents, and it is inspired by the work of Mitri [25, 26, 27]. In these works, the authors evolve populations of artificial agents in an arena containing a food and a poison feeder. Agents evolved to signal when they are in the vicinity of the food feeder. However, a setback of this work is their use of cloned individuals, as this behaviour is mostly found mainly when the members of the colony have high relatedness. Also, it is important to note that the main interest of the research and analysis reported in [25] lays on the agents signaling (emitters).

The work reported here, focuses on the agent perceiving the signal (perceivers), this is, agents exploiting the existence of a signal for their own benefit. The next significant difference in this work lays on the analysis of the possible behavioural effects due to a changing in the position of the sources of food and poison during the evolutionary process (see
Performance tests).

Three different families understood as the result of three different evolutionary processes are obtained and subjected to different tests. The paper is organized as follows: Section 2 describes de environment, robots and neuro-controllers used in the experiments. Section 3 describes the experiments and presents the results. Section 4 concludes the paper.

2 Environment and robot description

We used a simulated version of the e-puck robot [28], which is a two-wheeled mobile robot with differential motors, a blue LED ring, a 360° camera with a resolution of 720×1 pixels, and a ground camera capable of detecting the color of the floor. The e-puck robot is a suitable agent for research on inter-robot basic communication. The experiments were performed using the open-source physics-based robot simulator Enki [29] in an 2×2 m^2 arena delimited by a wall. Following [25], this arena contained two feeders placed on the floor and emitted red light, and the floor around the feeders was coloured (Figure 2. *The arena used for the Fixed location condition. Feeders emitting red light located at the opposite corners of the arena. Food: black circle surrounding the feeder at the top right corner. Poison: gray circle surrounding the feeder at the bottom left corner. The small empty circles represent some of the possible initial positions of the individuals.*. A black ring surrounded the feeder containing food, and a gray ring surrounded the feeder containing poison.

Neural controller

The behaviour of each robot was controlled by a feed-forward fully connected neural network. This neural controller had two layers, where there were eleven neurons in the input layer and three neurons in the output layer, making a total of the 33 connection weights (Figure 1. The neural controller used in all the experiments.). A linear activation function was used for each of the input neurons. Two of the input neurons were connected to the ground camera; one was activated when detecting food (black) and the other when detecting poison (gray). If the ground camera detected any of these sources, the input of the corresponding neuron was set to 1; otherwise, it was set to 0. Eight of the input neurons were used to encode the input image of the 360º camera, which was divided into four sections of 180 pixels each. For each section, the average of the blue and red channels was calculated and normalized within the range of 0 and 1. Therefore, four input neurons were used for the red values and four for the blue values. The last input neuron was established as a bias neuron with an input of 1 in any condition. The bias neuron was included with the intention of allowing the robots to produce some behaviour in the absence of sensory data.

The neurons in the output layer received the sum of each output from the input layer, multiplied by the weight of the corresponding connection. Two of the three output neurons were used for independently control the speed of the two wheels of the robot. The third output neuron was used to determine whether the robot emitted blue light or not. If the neuron's output value was higher than 0, the e-puck light ring was turned on. For every output neuron, the hyperbolic tangent function (1) in was used as the activation function.

$$f(x) = \frac{(e^{x}-e^{-x})}{(e^{x}+e^{-x})} \tag{1}$$

The values evolved with the genetic algorithm were the weights of the neural network connections. The weights were coded in a string of 264 bits, 8 bits for each of the 33 connections, allowing a range of 256 values per connection in an interval of -1 to 1.

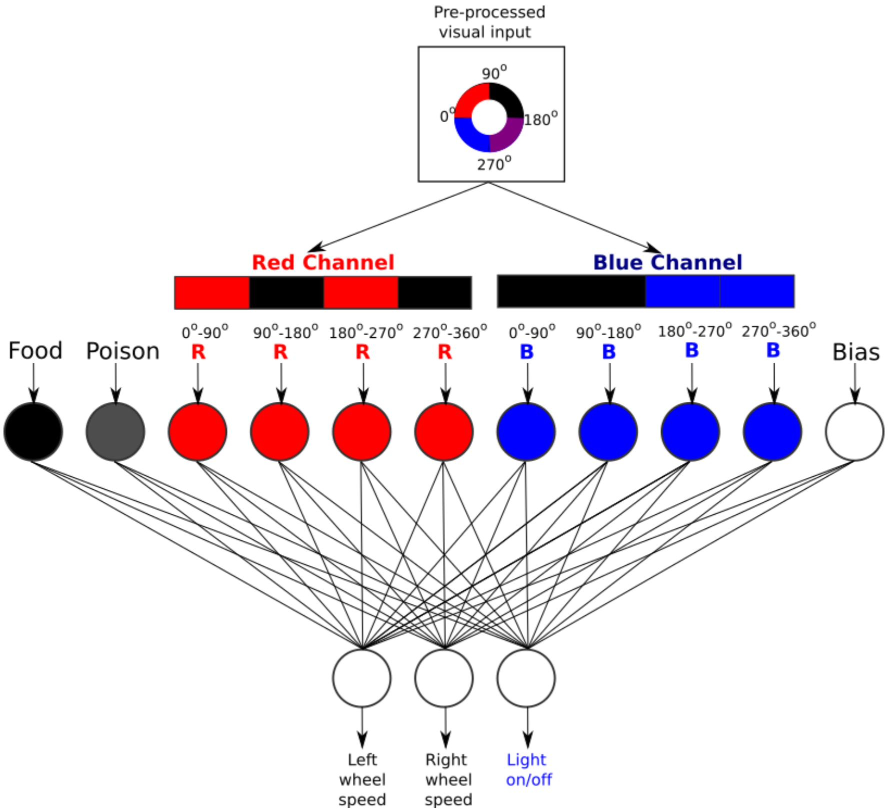


Figure 1. The neural controller used in all the experiments.

Evolutionary algorithm

In all the experimental conditions presented in this work, 100 genomes (individuals) were used as the starting population. Each genome coded 33 connection weights as a 264-bit string. The robots lived for 1200 cycles in groups of 10. Then, after being classified according to their fitness, 50 genomes were selected using the roulette wheel method. The 50 selected genomes were randomly paired. As in [30], with a 0.8 probability, these paired genomes were cloned and with a 0.2 probability were reproduced using two random crossover points. As a result, a new pool of 50 genomes was obtained and cloned to get the new 100 individuals for the next generation. All the individuals were first subjected to a genetic mutation, where each bit had a 0.01 probability of being mutated. The process was repeated for 250 generations.

3 Conditions, Experiments, and TESTS

The first condition aimed to obtain a basic communication behaviour between artificial agents with **fixed source location.** Following [30], the evolutionary runs were performed in an arena where the feeders (food and poison) were located in a fixed position in an arena. Two experiments under **fixed source location condition** were performed. In the first one, individuals were evolved to obtain a foraging behaviour (**forager individuals experiment**). The forager individuals experiment was designed to obtain a behaviour in the robots where it will be beneficial in terms of fitness to avoid the poison area, locate the food area and remain in it. In the second experiment, individuals evolved a signaling behaviour (**signalers individuals experiment**). The signaling individuals experiment was designed to develop a behaviour where emitting light on the food area will have a positive impact in terms of fitness.

Performance tests and **response to signaling tests** were made to verify that during the evolutionary process some type of basic communication strategy had evolved. The tests measure the response of evolved individuals to different sources of light. These tests were performed on a representative sample of the best individuals of each evolutionary algorithm (foragers and signalers). The performance tests aimed at measuring the total time these individuals spent in the regions of interest (ROIs: red light, blue light, and no light). The response to signaling tests were designed for making two relevant analysis. First, due that in these tests the location of the red light changed it allowed to remove any doubts concerning whether the agents only *learned* the location of the feeders in the arena on the fixed source condition. Second, these tests were designed to elucidate the effects of perceiving both lights combined in the same arena (blue-red vs red, blue-red vs blue tests). In other words, this response to signaling tests serves to elucidate the effect, the emitted blue light of other individuals together with the red light of the feeder, has on the behavioural preference of the individuals.

The second condition, **random source location**, was designed to test the effects of changing the location of the feeders during the evolutionary process. This condition is in line with the idea that in natural environments it is very unlikely to find food in the same location. In fact, biological agents must deal with a constantly changing context where most of the time the food location is unknown. Under these circumstances, evolving basic communication strategies may be highly valuable to efficiently find food sources. Therefore, under this condition only the **signalers individuals experiment** was performed, but both **performance tests** and **response to signaling tests** were also made. To our knowledge, this is the first time a direct comparison is made between the effects of fixed and changing locations of food sources when evolving a signaling behaviour.

Fixed source location condition

In this condition, the contextual configuration of the environment was fixed for all of the runs of the evolutionary algorithm. The evolutionary algorithm was run in a 2×2 m^2 arena, where two feeders of 20 cm diameter were placed and located 50 cm in the opposite corners of the arena. One feeder contained food and the other poison and they were differentiated by a surrounding black 25 cm diameter circular area for food, and a gray circular area for poison (see Figure 2. *The arena used for the Fixed location condition. Feeders emitting red light located at the opposite corners of the arena. Food: black circle surrounding the feeder at the top right corner. Poison: gray circle surrounding the feeder at the bottom left corner. The small empty circles represent some of the possible initial positions of the individuals.*). Both feeders were constantly emitting red light.

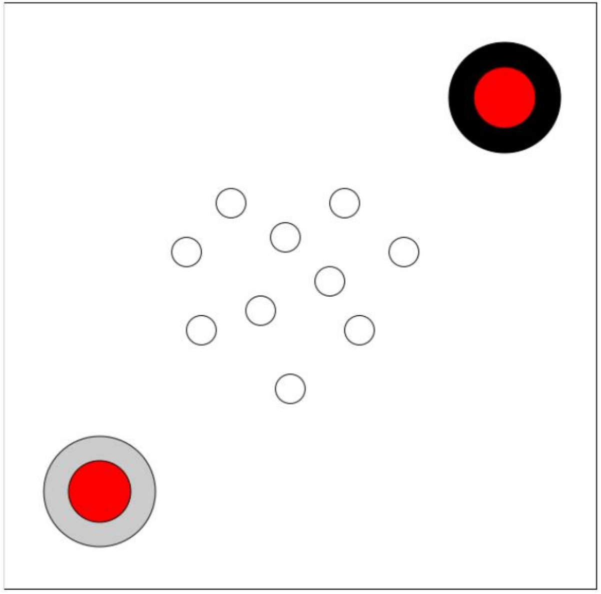


Figure 2. *The arena used for the Fixed location condition. Feeders emitting red light located at the opposite corners of the arena. Food: black circle surrounding the feeder at the top right corner. Poison: gray circle*

surrounding the feeder at the bottom left corner. The small empty circles represent some of the possible initial positions of the individuals.

Forager individuals experiment

This experiment aimed at obtaining robot controllers, which drive the agents to avoid the poison area and find the food area and then staying on it. It should be noted that even when the blue LED ring had no role in this task, it was still controlled by the neural network of the robot. The evolutionary processes were run 20 times, each starting with a random generation of 100 genomes. Each generation of individuals was divided into groups of 10 individuals, and each group was let to live for 1200 cycles.

At the beginning of all simulations, each group of 10 individuals was distributed at random positions in the environment (Figure 2), but with the restriction that the individuals cannot start the simulation within the areas of food or poison. Over the course of the simulation the fitness for each individual was calculated as the number of cycles that the individual spent on the food area (t_f) minus the number of cycles that the individual spent on the poison area (t_p) (2) :

$$\boldsymbol{F = \sum_{i=1}^{T} \left(t_f(i) - t_p(i) \right), \textit{if } F < 0 \textit{ then } F = 0} \tag{2}$$

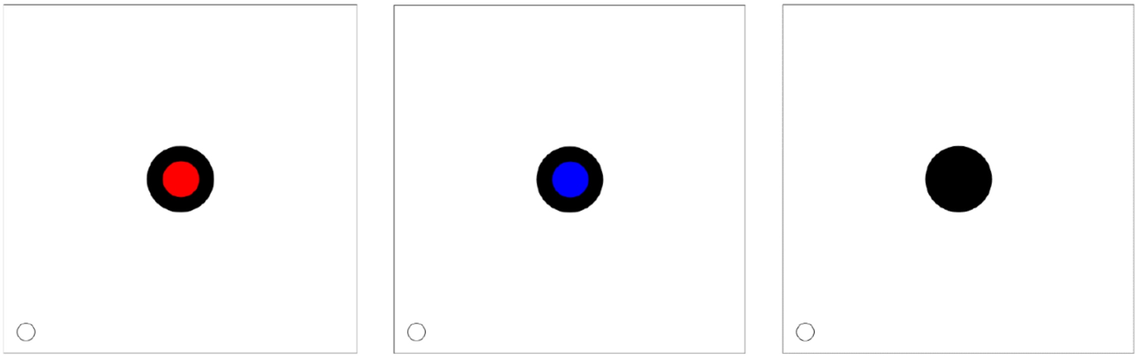
Although the blue light, which could be emitted by other agents had an effect on the output of the neural controller (when a value > 0), this was not coded in the fitness function. Therefore, the blue light emission, or the lack of it, did not play a role in the selection process and consequently on the evolutionary pressure. The evolutionary algorithm was run for 250 generations, and the process was repeated 20 times, to produce 20 final generations of 100 robots each. Hence, 20 final generations, $2{,}000$ individuals were obtained. Then, the performance tests and response to signaling tests were done on a sample of the best 600 individuals of each evolutionary algorithm. So, for each run, the best 30 individuals with the highest fitness scores were tested.

Performance tests

The performance tests aimed at measuring the total time individuals spent in the ROIs (red light, blue light, and no light). As explained before, these tests were designed to find what are the effects of perceiving the blue and red light sources, and the lack of it, on the robots' behaviours. To make these tests, the floor camera was occluded (the neural inputs from this camera were set to zero). Then, the total time an individual spent in each ROI during an epoch (1200 life cycles) was normalized to values between 0 and 1. The normalized total time individuals spent on each ROI was analyzed.

Three types of testing arenas were designed, each arena was of 2×2 m^2 and contained different ROIs (Figure 3). The ROIs were delimited with a circle of 20 cm diameter, surrounded by a black circle area of 25cm diameter. The initial position for all individuals was fixed at the left bottom corner of the arena. Of the total sample of 600 individuals, 200 individuals were randomly assigned to one of the three arenas: **red light arena**- a red light source was placed in the center of the ROI (Figure 3a); **blue light arena**- a blue light source was placed in the center of the ROI (Figure 3b); **no light arena**- no source of light, only a black circle was placed in the center of the ROI to measure the time spent on it (Figure 3c).

There were statistically significant differences between the group means of the total time spent in the three ROIs as determined by one-way Welch's ANOVA ($F(2, 279.247) = 224.397, p < 0.001$). A Games-Howell post hoc test revealed that the total time spent in the ROIs was significantly higher in the presence of red light ($M = 0.379 \pm SEM = 0.017, p < 0.001$), than in the presence of blue light ($M = 0.033 \pm SEM = 0.007, p = .002$), and to a greater extent when there was no light ($M = 0.005 \pm SEM = 0.001$). These results suggest that the controllers evolved to perceive red light as an indicator of food, and that the presence of blue light in the arena did not have any effect on the agents. Therefore, even though during the evolutionary process there were individuals emitting blue light, this did not affect the behaviour of the agents.



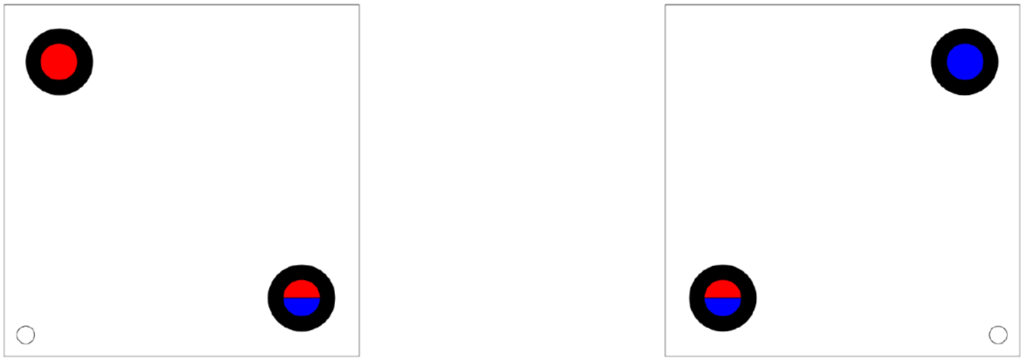
a) Red light arena b) Blue light arena c) No light arena

Figure 3. *The three arenas for the performance test. The black circle represents the ROI and the small empty circle the initial fixed position of the individuals.*

Response to signaling tests

The performance tests showed that the foragers individuals spent significantly more time in the red light ROI because it indicates food. However, it is highly probable that the individuals perceived, some amount of blue light combined with the red light. This situation could happen when agents surrounding the food feeder fortuitously switched the blue light on. Thus, the following tests were designed to measure the effects of this possible scenario by letting the agents perceive red and blue lights in the arena, and adding a ROI with blue and red light combined.

All 600 individuals were tested in 2 different arenas, one by one, for 1200 cycles. The time each individual spent in the ROIs was normalized. As in the performance tests, each arena was of $2 \times 2\ \mathrm{m}^2$, the ROIs had the same size, but here, two ROIs were located on the opposing corners of both arenas. The two arenas had the following configuration: **blue-red vs red arena**- one ROI emitting blue and red light combined, and a second ROI emitting red light (Figure 4a); **blue-red vs blue arena-** one ROI emitting blue and red light combined, and a second ROI emitting blue light (Figure 4b). The initial position for all individuals was one of the two empty corners of each arena. The two possible initial positions were equidistant from the ROIs.



a) Blue-red vs red arena b) Blue-red vs blue arena

Figure 4. *Arenas used for the response to signaling test with their respective ROIs. The empty circle represents the initial position of the individuals.*

Repeated measures of the total time an individual spent in each ROI were obtained because there were two ROIs in each arena. A two-tailed paired-samples t-test revealed that on average, in the blue-red vs red arena, individuals spent significantly more time in the red light ROI ($M = 0.205 \pm SEM = 0.009$), than in the combined blue-red light ROI ($M = 0.109 \pm SEM = 0.008$), $t(599) = -6.239, p < 0.001, r = 0.247$. Conversely, in the blue-red vs blue arena, individuals spent significantly more time in the combined blue-red light ROI ($M = 0.204 \pm SEM = 0.009$), than in the blue light ROI ($M = 0.008 \pm SEM = 0.002$), determined by a two-tailed paired-samples t-test, $t(599) = 18.954, p < 0.001, r = 0.612$.

The results remove all doubts about foragers individuals only learning the position of the food source. In fact, the results suggest that they evolved to locate the red light indicating a source of food and remain in it. Additionally, the individuals also reacted to the presence of a blue and red light combined, but a single red light source was preferred when it was present. The presence of a single blue light source had almost no effect on the behaviour of the individuals. Therefore, response to signaling tests suggest that individuals reacted to the red light as a signal for food.

Signalers individuals experiment

The main difference between foragers individuals experiment and the present experiment is that here, the emission of blue light was an important component in the evolution of the individuals. When an individual turned on its blue LED ring while being on a feeder it was considered as a signaling behaviour. Thus, this behaviour was included into the fitness (3) function of an individual as follows:

$$\boldsymbol{F = \frac{\sum_{i=0}^{T}\left(\left(l_f(i)\right) - \left(l_p(i)\right) - \alpha\left(l_s(i)\right)\right)}{T}, if\ F < 0\ then\ F = 0} \qquad (3)$$

Where l_f is the number of life cycles an individual turned on the blue LED ring while being in the ROI with food. This should be interpreted as a behaviour of signaling for food. l_p is the number of cycles an individual turned on the blue LED ring while being in the ROI with poison. In this case, the signal should be interpreted as signaling for poison. Finally, l_s is a general cost for emitting a signal and α a constant set empirically to 0.04. A total of 20 evolutionary processes were run using the arena of the fixed context condition (Figure 2). The evolutionary parameters used for the signalers individuals experiment were exactly the same as the ones used in the previous experiment. Additionally, just as for the forager individuals experiment, the performance tests and response to signaling tests were done on a sample of the best 600 individuals.

Performance tests

From the total of the 600 selected individuals, 200 individuals were randomly assigned to one of the three testing arenas (red light, blue light, no light). The same parameters from the previous performance tests were used here. The normalized total time that individuals spent on each ROI was analyzed.

There were statistically significant differences between group means of the total time spent in the three ROIs as determined by one-way Welch's ANOVA $(F(2{,}274.416) = 115.220, p < 0.001)$. A Games-Howell post hoc test revealed that the total time spent in the ROI was significantly higher in the presence of red light $(M = 0.275 \pm SEM = 0.018, p < 0.001)$, than in the presence of blue light $(M = 0.046 \pm SEM = 0.01, p < .001)$, and to a greater extent when there was no light $(M = 0.006 \pm SEM = 0.001)$.

The results suggest that the individuals responded to red light as an indicator of food. The behavioural preference for the red light persisted even when emitting blue light while being in the ROI with food was included into the fitness function. This finding is of particular relevance because the blue light could also be considered as an indicator of food. Given that emitting light on the food area had a positive impact in terms of fitness, it is more likely that during the evolutionary process, the signalers individuals were exposed to a combination of blue and red light to a greater extent than the foragers individuals. Therefore, contrary to what was found in the foragers individuals experiment, the signalers individuals should respond to a combined blue and red light source as signaling food. The following tests elucidated this hypothesis.

Response to signaling tests

The same sample of 600 signalers individuals was tested on the same two arenas as the foragers individuals (Figure 4). As expected, the results of these tests were the opposite as those found in the foragers individuals experiment. On average, in the blue-red light vs red light arena the signalers individuals spent significantly more time in the blue-red light ROI $(M = 0.134 \pm SEM = 0.008)$ than in the red light ROI $(M = 0.089 \pm SEM = 0.007), t(599) = 3.586, p < 0.001, r = 0.144$. Likewise, in the blue-red light vs blue light arena the signalers individuals spent significantly more time in the blue-red light ROI $(M = 0.183 \pm$

$SEM = 0.009$), than in the blue light ROI ($M = 0.007 \pm SEM = 0.002$), determined by a two-tailed paired-samples t-test, $t(599) = 18.204, p < 0.001, r = 0.596$.

These findings suggest that the presence of a combined blue and red light source had a greater influence on the behaviour of the individuals than a single red light source. Therefore, the signalers individuals evolved to respond to a greater extent to the red light of the food feeder together with the blue light emitted by other individuals when signaling food.

Random source location condition

The random location condition was inspired in how food sources constantly change in natural contexts. Evolving basic communication strategies may be highly useful to influence the behaviour of another individual to efficiently find food sources. Therefore, only the **signalers individuals experiment** was performed in the random location condition, making use of the (3) and all other variables of the process also being the same.
The location of the feeders in the arena was randomly selected from a pool of 32 predefined arenas. The distance between feeders varies in the range of 1.4 to 1.8 m. All of the arenas had a square shape of 4 m^2, and contained two circular ROIs of 20 cm diameter with a red light source at its center. Again, a black circle of 25cm diameter around the ROI represented food, and gray circle represented poison. As in the previous experiments the best 30 robots from each final generation of 20 runs were used for testing. The individuals were subjected to the same two tests.

Performance tests

From the sample of the 600 selected individuals, 200 were randomly assigned to each one of the three testing arenas (red light, blue light, and no light), and the same parameters of the previous performance tests were used. The normalized total time that individuals spent on each ROI was analyzed.

There were statistically significant differences between group means in the total time spent in the ROI of the testing arenas as determined by one-way Welch's ANOVA ($F(2{,}272.094) = 98.836, p < 0.001$). A Games-Howell post hoc test revealed that the total time spent in the ROI was significantly higher in the red light arena ($M = 0.245 \pm SEM = 0.018, p < 0.001$) than in the blue light arena ($M = 0.073 \pm SEM = 0.011, p < .001$), and in the no light arena ($M = 0.004 \pm SEM = 0.001$). Again, these results suggest that individuals responded to red light as an indicator of food and that the blue light by itself had an insignificant influence on individuals in their search for food. However, signalers individuals should respond to a greater extent to a combination of blue and red light source as an indicator of food. The following tests clarified this hypothesis.

Response to signaling tests

The same sample of 600 individuals used in the performance test were set in the arenas shown in Figure 4a and Figure 4b for 1200 cycles. The normalized total time individuals spent on both ROIs of each arena was analyzed. As expected, in the blue-red light vs red light testing arena individuals spent significantly more time in the blue-red light ROI ($M = 0.154 \pm SEM = 0.008$) than in the red light ROI ($M = 0.112 \pm SEM = 0.008$), $t(599) = 2.972, p = 0.003, r = 0.12$. Also, individuals spent significantly more time in the blue-red light ROI ($M = 0.209 \pm SEM = 0.009$) than in the blue light ROI ($M = 0.011 \pm SEM = 0.002$), $t(599) = 18.941, p < 0.001, r = 0.612$.

In general, the signalers individuals responded to a single red light source as a signal of food and ignored a single blue light source. Importantly, as in the fixed context condition, in the random context condition the signalers individuals responded to the combined blue and red light source as the one being the most reliable signal of food. However, it is still unclear in what way, or if in some way, individuals who evolved in the random context condition responded differently to this light source than those who evolved in a fixed context. The following analyzes were made to elucidate this question.

Fixed versus random location

To elucidate the effects of evolving in a fixed versus a random location of the feeders, the data from the 600 signalers individuals from both conditions were compared regarding the total time spent in the blue-red light ROIs. First, in the blue-red light vs red light testing arena, a two-tailed independent-samples t-test revealed that there were non-significant differences in the average time spent in the blue-red light ROI by individuals evolved in the fixed context $(M = 0.134 \pm SEM = 0.008)$ compared to the individuals evolved in the random context $(M = 0.154 \pm SEM = 0.008)$, $t(1195.915) = -1.637, p = 0.102$.

Second, in the blue-red light vs blue light testing arena, although there were non-significant differences, a slight tendency to spent more time in the blue-red light ROI was observed in individuals evolved in the random context $(M = 0.209 \pm SEM = 0.009)$ compared to the individuals evolved in the fixed context $(M = 0.183 \pm SEM = 0.009)$, $t(1194.875) = -1.946, p = 0.052$. This finding only indicates that evolving in a random context gives a slight evolutionary advantage to the signal-receiving individuals to differentiate food sources when there are other irrelevant sources of blue light in the same context.

These results were surprising, given that the signalers individuals evolved in the random location condition should rely to a greater extent on the blue light in combination with red light to efficiently find food sources. However, making a direct comparison between individuals evolved in fixed and random location may guide further research to a better understanding of how animal signaling evolves under changing conditions.

4 Discussion

It has been suggested that signals evolve from behaviours that were originally only cues [8, 31, 32]. That is, it can be expected that communication systems originate through inadvertently produced information that evolves into a signal [33].

The work presented in this article addresses, from an evolutionary approach, how a basic communication strategy can emerge in artificial agents through the co-evolution between signalers and receivers. With this approach, it was explored how different evolutionary conditions can determine the success of evolving signals. For the forager individuals, the aim of the task laid in locating the source of food and avoiding the source of poison, therefore, the action of signaling or not signaling did not directly affect the fitness of individuals. According to the results obtained in the performance and response to signaling tests, the impact of the response to the blue light was less than the response to the red light emitted by the feeders. This is due to the fact that, the fortuitous emission of blue light resulted to be less significant concerning the light emitted continuously by the feeders.

For the signalers individuals, in both conditions (fixed and random), signaling had an impact on the fitness of the individuals. Fitness was increased when signaling the source of food and decreased when signaling the source of poison and/or any other place. Although signaling implied a cost, given the positive impact of signaling for food, the robots were able to signal the food area as long as possible. While the results obtained in the performance test showed that the robots developed a greater preference for red light than blue light, the response to signaling test showed that a significant sample of the robots developed an attraction to blue/red light. This suggests that this attraction arose through information inadvertently produced by other robots by signaling the food source, which then evolved into a signal [34].

In comparison with similar studies [30] [25] [35], the contribution of this work focuses on how the co-evolution of signalers and receivers within small genetically variable populations can lead to the evolution of a basic communication strategy. This was achieved through the stochasticity of genetic operators, without the need of using *"extra"* processes that would lead to cloned individuals or individuals with specific levels of relatedness and/or selection.

An important aspect of the work presented here is the emergent property of attraction to blue light without this being a directly coded behaviour. In other words, the experimental results suggest that this attraction arose through information inadvertently produced by other robots by signaling the food source, which at the end evolved into a signal.

In the strict sense, every signal has a production cost that comes from the efficiency component even if it is negligible compared to other costs and benefits in the system [36, 37]. Consequently, an important attribute in the evolutionary process of signaling is the cost this has. Implementing a fitness function without a cost for signaling implies a greater space of solutions and a far greater number of non-optimal solutions. Given the experimental configuration in the experiments presented here, once an individual had located the source of food, the possibility of staying in its vicinity until the end of the test depended only on the phenotype expressed during its evolutionary process. Therefore, not penalizing the free signaling in the arena, would imply that fitness of the individuals was not affected by signaling an area other than food. To the best of our knowledge, this important detail had not been tested experimentally.

According to the results obtained in response to signaling test, the evolved controllers within a random source location proved to be more robust. This suggests that the randomness of the location of the feeders slightly improved the performance in solving the task. Having a different location in each evaluated generation and for each group of robots, lead to more generic controllers. Since, although the task was the same as in the fixed location condition, during the evolutionary process the robots had the possibility of exploring a greater number of solutions of the same task.

When evolved in a random source location, a slight tendency of the signalers individuals to ignore to a greater extent the blue light in the absence of red light was observed. However, there were no significant differences between the behaviour of signalers individuals evolved in a fixed and random source location condition. These results are in line with the role of functional information on the perceivers' behaviour. Here is suggested that, regardless of its location, the functional information of the signal remained constant for making adaptive decisions. Thus, the signal evolved in a fixed and in a random source location as an *egocentric spatial reference frame* for food [38]. Independently of the external location, the signal functional information referenced food location in a spatial relationship to the perceiver’s body.

Finally, in the known literature, performance is measured by the position of the moving individuals in relation to its original position (attraction and repulsion) after certain number of cycles, placing each individual at a specific distance from a second stationary individual that emits blue light [39, 25]. However, in this work, measuring the performance and the reliability of the communication strategy through the life cycles that individuals spent in the region of interest under different conditions, allowed the analysis of the performance of the perceivers and the task to be solved under different visual inputs (red, blue and red/blue light). In other words, and very importantly, individuals were evolved for signaling, however they were tested as perceivers.

The arguments by Carazo and Font [9] regarding the role of functional information are suitable to explain the present results. During the evolutionary process, signalers individuals do not aim to provide any information to other individuals in the arena. However, the emission of blue light in the presence of food made the influence of a signal adaptive for individuals when tested to perceive the food location. The best signaling individuals, those with the highest fitness, were tested when perceiving red light, blue light, and a combined blue and red light source. These individuals responded to the signal for food; the combined blue and red light function as a reliable source of food.

Carozo and Font [9] suggested that the fitness of signalers is in a direct relation to the effect or effects of a signal on the perceiver. The present work gives empirical evidence to this assumption. During the evolutionary process, signaling had an impact on the fitness of the individuals, and a direct relation to the influence on individuals when perceiving the signal was observed. When perceiving, signalers individuals made an ‘adaptive decision’ to spend more time in the combined blue and red source of light rather than in the single blue or red light source. Future work should better clarify the direct relation of these ‘adaptive decisions’ on the perceivers' fitness. A very interesting question to address is whether there is a correlation between a high fitness as signaler and the performance of the individual as a perceiver.

5 References

[1] J. A. Endler, "Some General Comments on the Evolution and Design of Animal Communication Systems," Philosophical Transactions of the Royal Society B: Biological Sciences, vol. 340, pp. 215-225, may 1993.
[2] R. B. Cocroft and R. L. Rodríguez, "The behavioral ecology of insect vibrational communication," BioScience, vol. 55, pp. 323-334, 2005.
[3] M. J. Owren, D. Rendall and M. J. Ryan, "Redefining animal signaling: influence versus information in communication," Biology & Philosophy, vol. 25, no. 5, pp. 755-780, 2010.
[4] D. Kalkman, "New problems for defining animal communication in informational terms," Synthese, vol. 196, no. 8, pp. 3319-3336, 2019.
[5] R. Dawkins and J. R. Krebs, "Animal signals: information or manipulation," Behavioural ecology: An evolutionary approach, vol. 2, pp. 282-309, 1978.
[6] J. M. Smith and D. Harper, Animal signals, Oxford University Press, 2003.
[7] Scott-Phillips and C. Thomas, "Defining biological communication," Journal of evolutionary biology, vol. 21, no. 2, pp. 387-395, 2008.
[8] J. M. Smith and D. Harper, Animal signals, Oxford University Press, 2003.
[9] P. Carazo and E. Font, "Putting information back into biological communication," Journal of evolutionary biology, vol. 234, no. 4, pp. 661-669, 2010.
[10] R. M. Seyfarth and D. L. Cheney, "Signalers and receivers in animal communication," Annual review of psychology, vol. 54, no. 1, pp. 145-173, 2003.
[11] A. V. Padje, M. D. Whiteside and E. T. Kiers, "ScienceDirect Signals and cues in the evolution of plant -- microbe communication," Current Opinion in Plant Biology, vol. 32, pp. 47-52, 2016.
[12] W. A. Searcy and S. Nowicki, The Evolution Of Animal Communication: Reliability And Deception In Signaling Systems, vol. 39, Princeton Univ Pr, 2005, p. 270.
[13] D. Cliff, P. Husbands and I. Harvey, "Explorations in Evolutionary Robotics," Adaptive Behavior, vol. 2, pp. 73-110, 1993.
[14] J. Lehman, J. Clune, D. Misevic, C. Adami, L. Altenberg, J. Beaulieu, P. J. Bentley, S. Bernard, G. Beslon, D. M. Bryson and others, "The surprising creativity of digital evolution: A collection of anecdotes from the evolutionary computation and artificial life research communities".
[15] R. E. Lenski, M. R. Rose, S. C. Simpson and S. C. Tadler, "Long-Term Experimental Evolution in Escherichia coli. I. Adaptation and Divergence During 2,000 Generations," The American Naturalist, vol. 138, pp. 1315-1341, dec 1991.
[16] J. Vonk and C. Aradhye, "Evolution of Cognition," Basics in Human Evolution, pp. 479-491, 2015.
[17] H. Lipson and J. B. Pollack, "Automatic design and manufacture of robotic lifeforms," Nature, vol. 406, no. 6799, pp. 974-978, 2000.
[18] S. Kriegman, N. Cheney and J. Bongard, "How morphological development can guide evolution," Scientific reports, vol. 8, no. 1, pp. 1-10, 2018.
[19] S. riegman, D. Blackiston, M. Levin and J. Bongard, "A scalable pipeline for designing reconfigurable organisms.," Proceedings of the National Academy of Sciences, vol. 117, no. 4, pp. 1853-1859, 2020.
[20] E. Coutinho, E. R. Miranda and A. Cangelosi, "Towards a model for embodied emotions," in Artificial intelligence, 2005. epia 2005. portuguese conference on, 2005.
[21] O. Gigliotta, G. Pezzulo and S. Nolfi, "Evolution of a predictive internal model in an embodied and situated agent," theory in biosciences, vol. 130, pp. 259-276, 2011.
[22] A. Bernard, A. Jean-Baptiste and N. Bredeche, "To cooperate or not to cooperate: why behavioural mechanisms matter," PLoS computational biology, vol. 12, no. 5, p. e1004886, 2016.
[23] P. Grouchy, G. M. D'Eleuterio, M. H. Christiansen and H. Lipson, "On the evolutionary origin of symbolic communication," Scientific reports, vol. 6, no. 1, pp. 1-9, 2016.
[24] D. Kadish, S. Risi and L. Beloff, "An artificial life approach to studying niche differentiation in soundscape ecology," in Artificial Life Conference Proceedings, MIT Press, 2019, pp. 52-59.
[25] S. Mitri, D. Floreano and L. Keller, "Relatedness influences signal reliability in evolving robots," Proceedings of the Royal Society of London B: Biological Sciences, p. rspb20101407, 2010.
[26] D. Floreano, Mitri, Sara, S. Magnenat and L. Keller, "Evolutionary Conditions for the Emergence of Communication in Robots," Current Biology, vol. 17, no. 6, pp. 514-519, 2007.

[27] D. Floreano, S. Mitri, A. Perez-Uribe and L. Keller, "Evolution of Altruistic Robots," in Computational Intelligence: Research Frontiers, Berlin, Heidelberg, Springer Berlin Heidelberg, 2008, pp. 232-248.

[28] F. Mondada, M. Bonani, X. Raemy, J. Pugh, C. Cianci, A. Klaptocz, S. Magnenat, J.-C. Zufferey, D. Floreano and A. Martinoli, "The e-puck, a robot designed for education in engineering," Proceedings of the 9th conference on autonomous robot systems and competitions, vol. 1, pp. 59--65, 2009.

[29] S. Magnenat, M. Waibel and A. Beyeler, "Enki: The fast 2D robot simulator," URL http://lis. epfl. ch/resources/enki, 2009.

[30] S. Mitri, D. Floreano and L. Keller, "The evolution of information suppression in communicating robots with conflicting interests," Proceedings of the National Academy of Sciences, vol. 106, pp. 15786-15790, 2009.

[31] T. C. Scott-Phillips, R. A. Blythe, A. Gardner and S. A. West, "How do communication systems emerge?," Proceedings of the Royal Society of London B: Biological Sciences, vol. 279, pp. 1943-1949, 2012.

[32] M. Fröhlich and P. van Schaik and Carel, "Must all signals be evolved? A proposal for a new classification of communicative acts," Wiley Interdisciplinary Reviews: Cognitive Science, p. e1527, 2020.

[33] M. Dorigo, M. Birattari, C. Blum, A. L. Christensen, A. Reina and V. Trianni, "Preface," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11172 LNCS, p. vi, 2018.

[34] R. Palacios-leyva, F. Aldana-franco, B. Lara-guzmán and F. Montes, "The impact of population composition for cooperation emergence in evolutionary robotics," International Journal of Combinatorial Optimization Problems and Informatics, vol. 8, pp. 20-32, 2017.

[35] D. Floreano, S. Mitri, S. Magnenat and L. Keller, "Evolutionary conditions for the emergence of communication in robots," Current biology, vol. 17, pp. 514-519, 2007.

[36] S. Számadó, "The cost of honesty and the fallacy of the handicap principle," Animal Behaviour, vol. 81, pp. 3-10, 2011.

[37] M. E. Laidre and R. A. Johnstone, "Animal signals," Current Biology, vol. 23, pp. R829--R833, 2013.

[38] F. Filimon, "Are all spatial reference frames egocentric? Reinterpreting evidence for allocentric, object-centered, or world-centered reference frames," Frontiers in human neuroscience, vol. 9, p. 648, 2015.

[39] S. Mitri, D. Floreano and L. Keller, "The evolution of information suppression in communicating robots with conflicting interests," Proceedings of the National Academy of Sciences, vol. 106, pp. 15786-15790, 2009.