INTELIGENCIA ARTIFICIAL

Augmenting Scalable Communication-Based Role Allocation for a Three-Role Task

Gustavo Martins[1,3], Paulo Urbano[1,A]

[1]BioISI, Faculdade de Ciências da Universidade de Lisboa, Lisbon, Portugal.
[A]pub@di.fc.ul.pt
[3]Instituto de Telecomunicações, Lisbon, Portugal

Abstract In evolutionary robotics role allocation studies, it is common that the role assumed by each robot is strongly associated with specific local conditions, which may compromise scalability and robustness because of the dependency on those conditions. To increase scalability, communication has been proposed as a means for robots to exchange signals that represent roles. This idea was successfully applied to evolve communication-based role allocation for a two-role task, with one communication channel. However, it was necessary to reward signal differentiation in the fitness function, which is a serious limitation as it does not generalize to tasks where the number of roles is unknown a priori. We show that rewarding signal differentiation is not necessary to evolve communication-based role allocation strategies for the referred two-role task, and we improve reported scalability, while requiring less a priori knowledge. We extend the previous work to a three-role task and we propose and compare two cognitive architectures, to increase the number of communication channels, and several fitness functions to evolve scalable controllers. Our results suggest that communication might be useful to evolve role allocation strategies for increasingly complex tasks.

Keywords: role allocation, collective robotics, swarm, NEAT, artificial evolution, evolutionary robotics.

1   Introduction

Research on evolutionary collective robotics [2] [17] [20] for homogeneous robots suggest that role differentiation, which is fundamental for cooperation in natural and artificial systems, is triggered by differences in local physical interactions. In [2], a group of four robots was evolved to collectively navigate toward a light. In the most successful strategy, the robot in the front right position assumed the guide role, setting the path, while the robots in other positions followed. This strategy clearly depended on the specific positions of the robots in the group. In [17], a team of three robots was evolved for the ability to navigate as a group. The most successful strategy relied on two phases: (i) robots negotiate their positions until they reach a line formation; (ii) the first robot moves backwards while the others move forward following the first robot. If any robot is removed, the remaining robots cease motion. In [20], a group of five robots was evolved to guard a nest and forage simultaneously. The environment had two variations that required different behaviors to maximize fitness: (i) most robots stayed in the nest while others foraged; or (ii) fewer robots stayed in the nest while others foraged. Each variation had a corresponding nest color that robots could detect. The nest was divided into six sectors and each robot was randomly placed in a sector. The study concluded that the role each robot assumed depended on the nest color and the robot's position at the nest. For the above studies, the role assumed by a robot depends directly on very specific local conditions potentially decreasing robustness and scalability of the evolved strategies.

To evolve scalable and robust solutions for role allocation, one study [8] proposed endowing robots with one communication channel, so that role allocation might be negotiated through the exchange of signals. Communication allowed robots to emit a signal with a numeric value in the range [0, 1] and the goal was to have one robot emitting a high value, and the others a low value. Although scalable solutions were evolved, there was no actual behavioral task for the robots to execute. Another study [9] further extends this work by introducing a two-role, double patrolling task (Figure 1), which we replicate, and further develop into a three-role task in our study.
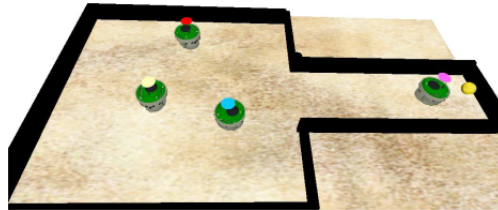

Figure 1: The double patrolling task environment. One robot must enter the corridor to find a light source while all others keep away. Reprinted with permission.

To evolve suitable controllers for the double patrolling task, Gigliota et al. [9] rewarded signal differentiation in the fitness function. This reward constitutes a priori knowledge, which is knowledge that is included by the system designer, to favor the evolution of a desired behavior but also constrain the evolutionary process [10]. Including less a priori knowledge imposes fewer restrictions on evolution, which can lead to more adaptive behaviors [14]. Furthermore, the authors [9] used a fixed-topology neuroevolution algorithm. However, different topologies result in different evolved behaviors and scalability for the same task [7]. To avoid the shortcomings of manual design [18], we substitute the fixed-topology neuroevolution algorithm by NEAT [19], a well known and widely applied, neuroevolution algorithm, that combines the search for appropriate network weights with a search for appropriate network topology.

Our goal is to evolve controllers with less dependency on a priori knowledge than required in previous research [9]. We aim to attain more robust and general solutions for the evolution of communication-based role allocation. We replicate the double patrolling task, and conduct experiments with a novel fitness function that does not reward signal differentiation. To determine the impact of the number of robots used during evolution on the scalability of the evolved solutions, we vary the number of robots used during evolution. Although communication-based role allocation was evolved in [9], the relevance of communication for performing the task and for evolving scalable strategies was not fully determined, because no control experiment without communication ability was conducted. To assess the relevance of communication for the task, we conduct experiments with no communication ability. We follow our previous work [11] and introduce a more challenging, three-role task, where two lights exist and which requires more than one communication channel. The goal is to have exactly one robot at each light while all other robots avoid approaching the lights. To increase the number of communication channels available to robots and evolve scalable controllers, we propose and conduct experiments with two cognitive architectures with two and three communication channels and five and six role sensors and actuators, while we also experiment fitness functions with and without reward for signal differentiation.

2   Related Work

Gigliota et al. [8] proposed the ability of robots to communicate their adopted roles by means of a numeric signal, to increase scalability of role allocation strategies. This approach facilitated the evolution of scalable strategies for the double patrolling task [9], where a group of robots has to allocate a sole robot to move to a light source while the other robots must stay away from light. In the most successful strategy, one robot emits a signal with a high value and assumes the role of exploring the environment to find the light while the other robots emit low values and avoid moving. Although evolution was conducted for groups of four robots, the evolved strategy was scalable for groups of 2-8 robots. The robots' controllers were neural networks evolved by means of an artificial evolutionary process. Artificial evolution depends

on, amongst other factors, a fitness function to measure the quality of the evolved behaviors. Fitness function design is known to be a critical aspect of evolutionary robotics [3], [16], [15], [6].

Two fitness functions were used in the referred study [9], BF and CRF. The first, BF (Eq. 1), rewards a closer distance between light and the closest robot, while rewarding a wider distance between light and all other robots.

$$BF = 0.75 \times BFC1 + 0.25 \times BFC2 \tag{1}$$

$BFC1$ (Eq. 2) rewards controllers able to have one robot close to light while $BFC2$ (Eq. 3) rewards controllers able to have all other robots away from light. These components are computed as follows:

$$BFC1 = \frac{1}{T} \times \sum_{t}^{T} \frac{\max(0, (M - d_t(L, light)))}{M} \tag{2}$$

$$BFC2 = \frac{1}{T \times (N - 1)} \times \sum_{t}^{T} \sum_{i}^{F} \frac{\min(M, d_t(F_i, light))}{M} \tag{3}$$

where $M = 0.9$ m is a maximal distance, $d_t(L, light)$ is the distance between light and its closest robot at instant t, T is the number of total time steps of a trial, F_i is robot i – excluding the closest robot to light, $d_t(F_i, light)$ is the distance between light and robot F_i at instant t and N is the number of robots.

The second fitness function, CRF (Eq. 5), extends BF to include a reward for signal differentiation, CFC: one robot emits a high signal and all other robots emit low signals.

$$CRF = 0.8 \times BF + 0.2 \times CFC \tag{4}$$

where CFC (Eq. 5) is computed as follows:

$$CFC = \frac{\sum_{t}^{T} \sum_{i}^{N} O_{t,max} - O_{t,i}}{T \times (N - 1)} \tag{5}$$

where $O_{t,max}$ is the highest value emitted at instant t, $O_{t,i}$ is the value emitted by robot i at instant t, and N is the number of robots in the group.

To illustrate how these functions behave, let us assume one robot at 0.0 m from light, a second robot inside the corridor at 0.25 m from light and all other robots in the home area at 0.75 m from light. One robot emits a 1.0 signal and all others emit 0.0. Table 1 shows fitness values according to BF and CRF, for such a situation, with different numbers of robots.

Table 1: BF and CRF fitness values for different numbers of robots.

Robots	BFC1	BFC2	BF	CFC	CRF
2	1	0.28	0.82	1	0.86
4	1	0.65	0.91	1	0.93
6	1	0.72	0.93	1	0.94
10	1	0.77	0.94	1	0.95

As the number of robots increases, fitness increases, with no performance improvement, because the second robot inside the corridor weights less in the group of robots supposed to be in the home area. The higher the number of robots, the lower the negative impact on fitness of a second robot in the corridor and the less effective BF becomes to evolve suitable controllers. Furthermore, BF is inadequate to measure performance when the number of robots in a group may vary, because variation in the number of robots has an impact in the fitness value, without any change in performance. The CRF fitness function shares the above limitations because BF is a component of CRF.

In [9], with BF the authors were not able to evolve communication use or any strategy to perform the task. Only with CRF were they able to evolve communication use and a scalable strategy, suggesting that communication is a potential solution to support scalability for role allocation strategies. However, CRF

rewards signal differentiation, which is undesirable because it forces a specific communicational scheme – one robot emitting a high signal and all other robots emitting low signals – hindering the system's ability to find other potentially suitable communicational schemes. We avoid such reward with the introduction of a novel fitness function presented in the next section (Eq. 6).

3   Two-Role Task

3.1   Evolving role allocation with less a priori knowledge

To evolve scalable strategies for the two-role task in [9], it was necessary to reward a signal differentiation scheme – one robot emits a high signal and all other robots emit low signals – in the fitness function, CRF. This approach, which resulted in the evolution of controllers able to perform the given two-role task, was possible because the authors were able to find, a priori, the above signal differentiation scheme which suits the two roles required for the task: one robot near the light and all other robots far away. However, as the number of roles needed to perform more complex tasks increases it is likely that the complexity of the required communicative behavior also increases. Therefore, it is unclear how rewarding signal differentiation could be used when a signal differentiation scheme that suits the roles required to perform the task can not be found a priori. As the task complexity increases, it becomes increasingly challenging to determine the specific signal differentiation scheme a priori to reward in the fitness function.

To avoid the reward for signal differentiation used in CRF and the dependency of BF on the number of robots, we introduce the TCD fitness function (Eq. 6). This function accounts for the existence of a sole robot in the corridor and the distance between light and the closest robot to light. TCD does not account for the number of robots as BF nor the communicative behavior as CRF.

$$TCD = \frac{1}{T} \times \sum_{t}^{T} \begin{cases} D_{t,light} & r = 1 \\ -D_{t,light} & r > 1 \\ 0 & r = 0 \end{cases} \tag{6}$$

where T is the number of time steps, r is the number of robots in the corridor (a robot is inside the corridor when its body center is inside the corridor) and $D_{t,light}$ is determined as shown in Equation 7.

$$D_{t,light} = \frac{\max(0, (Range - d_t(L, light)))}{Range} \tag{7}$$

where $Range$ is the light sensor range and $d_t(L, light)$ is the distance between light and its closest robot, at instant t. The closer to light this robot is, the higher $D_{t,light}$ is in range [0,1]. Fitness is $D_{t,light}$ if there is only one robot in the corridor and $-D_{t,light}$ when two or more robots are in the corridor. Otherwise, fitness is zero. To avoid negative fitness values, if the accumulated fitness up to instant t is less than zero, fitness is set to zero.

3.2   Experimental Setup

We use simulated e-puck [12] robots, that have a body diameter of 7.4 cm and distance between wheels of 5.2 cm. Robots have two independent wheel actuators to set the speed of each wheel, in [-0.1, 0.1] m/s, and a role actuator to emit a signal containing a decimal number in [0, 1]. Robots have eight obstacle sensors, equally spaced on the perimeter of the circular body, which measure the proximity of another robot or a wall, within 0.2 m; eight light sensors, also placed on the perimeter of the body, which measure the proximity to light, within 0.3 m; one non-directional role sensor with a range of 1.2 m which perceives the highest signal emitted by any other robot from any position; and a sensor to perceive the signal emitted by the robot itself in the previous time step. Each robot in the group is controlled by a copy of the same neural network and each input neuron receives values from one sensor and each output neuron sends values to one actuator. To simulate noise, each of these values is multiplied by a random number in range [0.95, 1.05]. Values coming from the sensors into the network are normalized in [0, 1] where closer proximity is represented by a higher value. Neural network output values are also normalized in [0, 1].

The experimental environment we replicated is composed of a 0.6x0.6 m area – the home – with an opening to a 0.2 m wide and 0.5 m long corridor (Figure 1). At the end of the corridor, there is a light source that robots cannot perceive from home. At the beginning of each experimental trial, robots emit a random signal and are placed in random positions and orientations in the home area, ensuring that no robots are colliding. Our experiments were run in the JBotEvolver [5] simulation platform.

In each experiment, we conducted 30 evolutionary runs of 1000 generations. The population has 100 individuals. In every generation, 15 trials with random initial conditions are generated and every individual in a generation faces the same set of trials. The fitness of an individual – a neural network – is the average fitness obtained in all those trials. A trial has a maximum duration of 2000 time steps. However, if a collision occurs the trial is terminated immediately to promote solutions that do not rely on or cause collisions. The NEAT implementation we used is Neat4J [13] with standard parameters.

In our experiments, we use the two fitness functions described above, BF (Eq. 1) and CRF (Eq. 4), as well as a new fitness function, TCD (Eq. 6). We conducted separate evolutions for groups of four robots, six robots, and ten and two robots where the numbers two and ten were randomly chosen by the simulator for each trial. In some experiments, we removed the robots' ability to communicate by removing the role sensor, as control – these experiments' names have the prefix "noRole".

3.3   Results

We post-evaluated the evolved controllers following the methodology used in [9] to allow for a direct comparison: only the controller that achieved the highest fitness during evolution is post-evaluated; the number of robots in the group varies between two and ten (we extended to twelve); each group with a different number of robots is post-evaluated in 100 trials; the post-evaluation function measures the percentage of time steps, within the last 100, when there is a sole robot in the corridor; satisfactory performance is one that shows a minimum post-evaluation fitness of 80%.

3.3.1   Evolution with Four Robots

In Figure 2, we show the post-evaluation results for groups evolved with four robots.
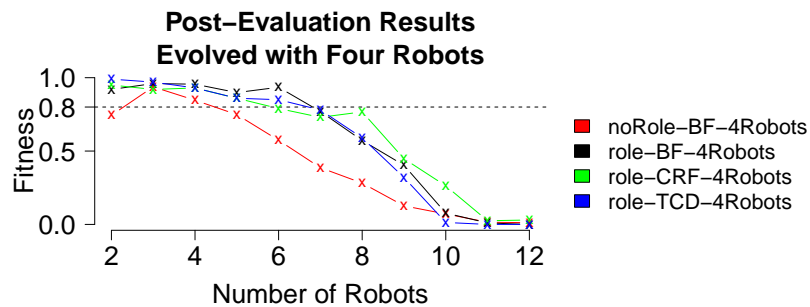

Figure 2: Post-Evaluation results when evolved with four robots

With no communication ability   For the noRole-BF-4Robots controller, performance is above 80% for groups of three and four robots, showing that communication is not strictly necessary to perform the task. Robots move in straight paths from corner to corner. To avoid collisions, robots change paths and eventually one robot enters the corridor and moves to the light. With less robots in the home area, there are less paths interfering and thus, it is less probable that another robot enters the corridor. If another robot enters the corridor, it moves towards the light, perceives another robot at the light and leaves the corridor. This strategy does not scale because the higher the number of robots, the higher the probability that interferences occur causing robots to enter the corridor.

With communication ability   Controllers role-BF-4Robots and role-TCD-4Robots, show a common strategy and a similar performance (Friedman $p = 0.179$), scaling for two to six robots. Robots move in

random directions emitting a 0.0 signal, until a robot finds the light and emits a 1.0 signal. All other robots continue to emit 0.0, leaving the corridor if inside, and change their motion pattern to small orbits in place, decreasing the probability of entering the corridor. However, a robot might enter the corridor to avoid a collision. In such case, the robot moves to the light, detects another robot nearby and leaves. The higher the number of robots, the more challenging it is for robots to maintain an orbital path in the home area while avoiding collisions, and the higher the probability of extra robots entering the corridor.

The role-CRF-4Robots controller scales for groups of two to five robots and shows a different strategy: before any robot enters the corridor one robot emits a 1.0 signal – the leader – and all other robots – the followers – emit a 0.0 signal. The leader moves along the wall while the followers orbit in place. If the leader detects another robot in the way, it relays the leadership to the detected robot and becomes a follower. Eventually, the leader enters the corridor and finds the light. Exceptionally, a follower enters the corridor to avoid a collision, moves to the light and becomes a leader. The previous leader, still in the home area, becomes a follower. A similar relay strategy was evolved in [7], with CRF, where different topologies where manually chosen and evaluated. Interestingly, for the other top controllers, a different strategy was evolved. Robots move in random directions in the home area; a robot enters the corridor to avoid a collision, finds the light and emits a 1.0 signal. The other robots maintain their behavior but avoid the corridor entrance. As the number of robots increase, the ability to avoid the corridor entrance decreases due to path interferences between robots.

The strategy evolved with the BF fitness function is scalable for 2-6 robots while in the previous work [9], no strategy that performs the task was evolved with BF. This improvement over previous work, where a fixed-topology neuroevolutionary algorithm was used, illustrates how a non fixed-topology neuroevolutionary algorithm may be more powerful when it comes to explore the solutions space. Furthermore, scalability is observed only in the communicative controllers, which suggest that communication is a relevant factor to the evolution of scalable role allocation strategies.

3.3.2 Evolution with Six Robots

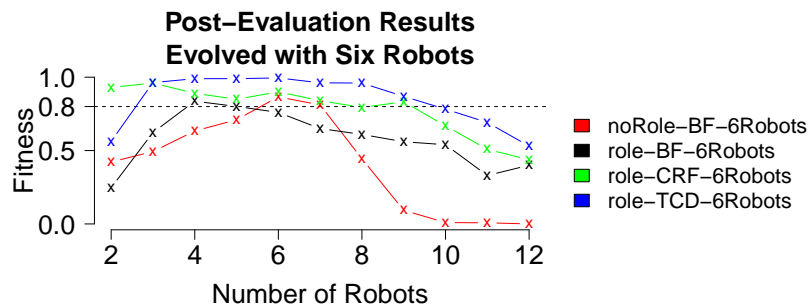In Figure 3, we show the post-evaluation results for groups evolved with six robots.

Figure 3: Post-Evaluation results when evolved with six robots

With no communication ability   Controller noRole-BF-6Robots shows satisfactory performance for six and seven robots but does not scale. Robots describe small elliptical paths in place, avoiding other robots within the obstacles sensor range. In this process, one of the robots enters the corridor and moves to the light. If another robot enters the corridor, it detects the first robot and leaves.

With communication ability   Controller role-BF-6Robots shows poor performance. When a robot finds the light and emits 1.0 all other robots also emit 1.0 and spin in place, even if that place is inside the corridor. This is a crucial difference to the previous communicative strategies evolved, where extra robots inside the corridor would leave. This strategy attained the highest fitness during evolution because the BF fitness function allows high fitness when more than one robot is inside the corridor, in spite of that behavior being the opposite of the desired.

For the role-CRF-6Robots and role-TCD-6Robots controllers, performance improved for seven, eight and nine robots, when compared to the 4-Robots experiments. For both controllers, the evolved strategy is similar to the main strategy described earlier: robots explore the environment until one robot finds the light and emits a 1.0 signal; if another robot is in the corridor, it leaves; robots in the home area change their behavior to avoid entering the corridor, but if another robot enters the corridor, it detects the robot at the light and leaves. Exceptionally, controller role-TCD-6Robots shows poor performance for two robots, because robots follow the walls instead of exploring in random directions. Robots cannot distinguish walls from fellow robots and when the group is composed of two robots distant from any wall, they follow each other in a circle as if they were following a wall, entering into a deadlock. Nevertheless, the strategy evolved with the TCD fitness function shows a higher post-evaluation fitness when compared to the strategy evolved with CRF (Friedman $p = 7.78 \times 10^{-5}$ for the null hypothesis of a similar performance for CRF and TCD).

3.3.3   Evolution with Two and Ten Robots

In Figure 4, we show the post-evaluation results for groups evolved with two and ten robots.
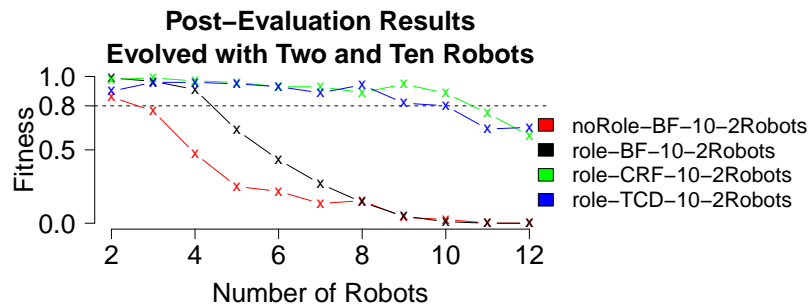

Figure 4: Post-Evaluation results when evolved with two and ten robots

With no communication ability   For the noRole-BF-10-2Robots controller, robots follow the wall, in the home area. The first robot entering the corridor, moves to the light. The other robots always enter the corridor but leave after detecting the robot at the light. This strategy works well for two robots because the second robot spends more time exploring the home area before re-entering the corridor than inside the corridor, thus minimizing the time an extra robot spends inside the corridor. As the number of robots increases, though, the time extra robots spend inside the corridor, increases as well, resulting in a lower post-evaluation fitness.

With communication ability   For the role-BF-10-2Robots controller, an alternative communication use has evolved. Robots explore the environment, emitting a 1.0 signal, instead of 0.0. When a robot moves to the light emits a 0.0 signal. If the group is composed of only two robots, the remaining robot in the home area changes the motion pattern to small orbits in place and the task is accomplished. However, if the group is composed of more than two robots, the robots in the home area do not change the motion pattern, for two reasons: (i) robots can only perceive the highest signal being emitted and, (ii) robots in the home area are emitting 1.0, the highest possible signal. Thus, the 0.0 signal being emitted by the robot at the light is not perceived by any robot. This instance illustrates how the desired solution was so deeply hardwired in the robots' design: by having a role sensor that only detects the highest signal being emitted, the system designers forced the robot at the light to use the highest possible signal, to inform it has found the light, in order to assure that the signal is perceived by the other robots. The communicational scheme evolved in this experiment was possible because the BF fitness function attains higher fitness for larger groups of robots, as illustrated in Table 1. In this experiment, the evolved communicational scheme attains high BF fitness for two and ten robots because it allows the task to be performed with precisely two robots and also because a group of ten robots is large enough to attain high fitness. In other words, according to BF, the evolved strategy is adequate for two robots and not

so inadequate for ten robots. However, post-evaluation results show that, after all, the behavior is not advantageous for a high number of robots.

For the role-CRF-10-2Robots controller, the evolved strategy is the main communicative strategy described earlier because the CRF fitness function forces the desired communicational scheme. For the role-TCD-10-2Robots controller, when the robots at the home area receive the 1.0 signal emitted by the robot at light, they spin in place and avoid the corridor, as seen before. Controllers role-TCD-10-2Robots and role-CRF-10-2Robots show no statistically significant difference in performance (Friedman $p = 0.062$) and the highest post-evaluation performance of all experiments, increasing previously reported scalability in [9] of 2-8 robots to 2-10 robots. TCD requires less a priori knowledge and is thus preferable over CRF.

4   Three-Role Task

For the three-role task experiments, we changed the environment, evolutionary setup and the sensors and actuators. To increase the number of roles to three, we introduced a light at the left side of the environment, at 1.40 m from the right light. We also removed the walls to decrease the duration of the experiments. At the start of a trial, robots are randomly placed at the center, between the lights. The goal is that exactly one robot moves to each light while the other robots avoid lights. Figure 5 shows one possible initial state and Figure 6 shows one desired final state. Numbers represent IDs.
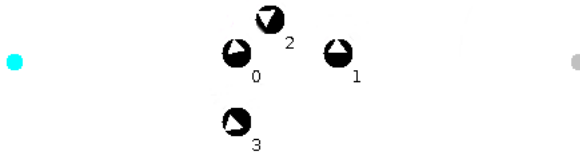

Figure 5: One possible initial position for the three-role task environment.

Figure 6: One desired final position for the three-role task environment.

On one hand, removing the walls decreases simulation time. On the other hand, it allows robots to wander off. To avoid robots wandering off, we increased the light sensors range to 2 m, allowing robots to use lights as reference points. An alternative to increase the light sensors range, would be to introduce an absolute reference point, and the corresponding sensor, which would increase the evolutionary search space and, thus, increase simulation time. Each robot has eight light sensors to detect the right light and eight light sensors to detect the left light. The role sensor range was increased to 2 m, because of the larger dimensions of the three-role task environment. The obstacle sensors range was decreased to 10 cm, as these sensors are used only to avoid collisions between robots and a larger range is unnecessary. The wheels actuators were maintained as in the two-role task. For each experiment, we conducted ten evolutionary runs, with 1500 generations, a population of 100 individuals, a trial duration of 400 steps and groups of four robots.

We experimented two architectures for increasing the number of channels: OneRoleAct, which relies on one role actuator and TwoRoleAct, which relies on two role actuators. In OneRoleAct, robots have one role actuator, one sensor to perceive the own role in the previous time step and three role sensors which are binary. Each role sensor is sensitive to a specific interval and outputs the value 1.0 when it detects a role actuator emitting in the interval it is sensitive to, or 0 otherwise. The first role sensor is sensitive to the interval $[0,\frac{1}{3}[$, the second sensor to $[\frac{1}{3},\frac{2}{3}[$ and the third sensor to $[\frac{2}{3},1]$. The three role sensors represent three channels. In TwoRoleAct, robots have two sets composed of one role actuator, one

role sensor which perceives the highest value being emitted by role actuators within range and one sensor to perceive the own role in the previous time step. The two sets represent two channels and the actuators and sensors of a channel are unable to exchange information with the actuators and sensors of the other channel. For each architecture, we conduct two sets of experiments, in a total of four experiments. In each set of experiments we use two fitness functions: one behavioral fitness function, that considers only the positional behavior of the robots, and a fitness function that also considers signal differentiation.

Behavioral Fitness Function for OneRoleAct and TwoRoleAct   The behavioral fitness function we used for OneRoleAct and TwoRoleAct, TCD_{lights}, shown in Equation 8, rewards a situation where one robot is close to a light, another robot is close to the other light while the other robots are away from any light (adapted from TCD, Eq. 6).

$$TCD_{lights} = \frac{1}{N} \times \sum_{n}^{N} TCD_{light_n} \tag{8}$$

where TCD_{lights} is the behavioral fitness for an environment with N lights and TCD_{light_n} is the behavioral fitness associated with light n, computed as shown in Equation 9.

$$TCD_{light_n} = \frac{1}{T} \times \sum_{t}^{T} \begin{cases} D_{t,light_n} & r = 1 \\ -D_{t,light_n} & r > 1 \\ 0 & r = 0 \end{cases} \tag{9}$$

where T is the number of time steps, r is the number of robots closer than 0.40 m to light n and $D_{t,light_n}$ is computed as shown in Equation 10.

$$D_{t,light_n} = \frac{\max(0, (K - d_t(L, light_n)))}{K} \tag{10}$$

where K is a maximum distance a robot may be from light n to increase fitness, arbitrarily defined as 0.30 m, and $d_t(L, light_n)$ is the distance between light n and its closest robot, at instant t. The closer to light n this robot is, the higher $D_{t,light_n}$ is in interval [0,1]. Fitness is $D_{t,light_n}$ if there is only one robot at a distance of at most 0.40 m from light n and $-D_{t,light_n}$ when two or more robots are closer than 0.40 m from light n. Otherwise, fitness is zero. To avoid negative fitness values, if the accumulated fitness up to instant t is less than zero, fitness is set to zero. We use the threshold 0.40 m because it was the value used for the two-role task, as it was the distance from the corridor entrance to light.

Signal Differentiation Reward Fitness Function for OneRoleAct   The fitness function that includes a reward for signal differentiation for the OneRoleAct experiment, $Comm_{OneRoleAct}$, shown in Equation 11, rewards a situation where one robot emits in $[0,\frac{1}{3}[$, another robot emits in $[\frac{2}{3},1]$ and all other robots in $[\frac{1}{3},\frac{2}{3}[$.

$$Comm_{OneRoleAct} = 0.75 \times BF_{lights} + 0.25 \times CS \tag{11}$$

where BF_{lights_n} is computed as shown above in Equation 8 and CS is the component that rewards signal differentiation, according to the number of robots emitting a signal on each channel, computed as shown in Equation 12.

$$CS = \frac{1}{T} \times \sum_{t}^{T} O_t + D_t + H_t \tag{12}$$

where T is the number of time steps, $O_t = 1/3$ when there is exactly one robot emitting in $[0, \frac{1}{3}[$, otherwise is zero, $D_t = 1/3$ when all robots except two are emitting in $[\frac{1}{3},\frac{2}{3}[$, otherwise is zero, and $H_t = 1/3$ when there is exactly one robot emitting in $[\frac{2}{3},1]$, otherwise is zero.

Signal Differentiation Reward Fitness Function for TwoRoleAct The fitness function that includes a reward for signal differentiation for the TwoRoleAct experiment, $Comm_{TwoRoleAct}$, shown in Equation 13, rewards a situation where one robot emits a high value on the first channel and a low value on the second channel, one robot emits a low value on the first channel and a high value on the second channel and all other robots emit low values on both channels.

$$Comm_{TwoRoleAct} = 0.75 \times BF_{lights} + 0.25 \times CM \qquad (13)$$

where BF_{lights} is computed as shown in Equation 8 and CM is the component that rewards signal differentiation, computed as shown in Equation 14.

$$CM = 0.75 \times \frac{1}{A} \times \sum_{a}^{A} SD_a + 0.25 \times \frac{1}{T} \times \sum_{t}^{T} R_t \qquad (14)$$

where A is the number of channels, T is the number of time steps and R is a reward for not having the same robot emitting the highest signal on both channels. At each time step, when the highest signals on all channels are emitted by different robots, R_t is set to one, otherwise R_t is set to zero. We introduced R to increase the probability of evolving solutions where the highest signal on each channel is emitted by different robots. SD_a is the signal differentiation reward for channel a, computed as shown in Equation 15.

$$SD_a = \frac{\sum_{t}^{T} \sum_{i}^{N} O_{at,max} - O_{at,i}}{T \times (N-1)} \qquad (15)$$

where $O_{at,max}$ is the highest value emitted at instant t, on channel a, $O_{at,i}$ is the value emitted by robot i, on channel a, at instant t, and N is the number of robots in the group.

4.1   Results

Figure 7 shows boxplots that represent the highest fitness controller for each run. The dotted lines represent the minimum fitness obtained when robots complete the task in an expeditious manner, i.e., after an initial phase with a duration of less than 80 steps, where role negotiation may happen, two robots move towards the lights while the other robots avoid the lights. Experiment names have the suffix "-Comm" when the signal differentiation reward is present and no suffix when the the reward is not present.
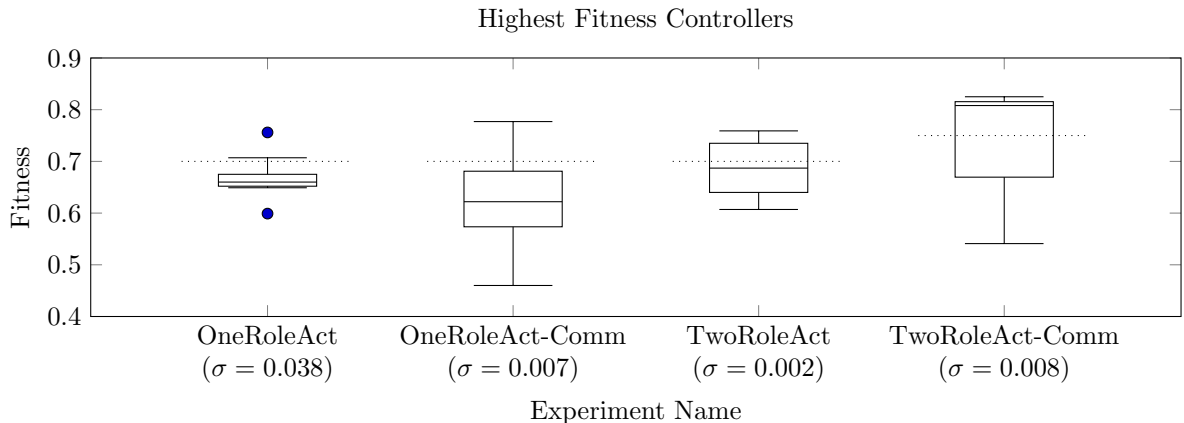

Figure 7: Boxplots for the highest performing controller for each run.

One controller able to perform above the minimum performance threshold evolved in all experiments. TwoRoleAct and TwoRoleAct-Comm evolved a higher number of controllers able to perform above the minimum threshold, when compared to OneRoleAct and OneRoleAct-Comm.

4.1.1   Post evaluation

We conducted 100 post evaluation trials for the best controller of each run. The post-evaluation function measures the percentage of time steps, within the last 100, when there is exactly one robot at a distance of 10 cm or less from each light and no other robots closer than 30 cm to any light. This three-role task post-evaluation function is more challenging than the two-role task post-evaluation function, as it requires robots to be closer to light – 10 cm instead of 50 cm. Satisfactory performance is one that shows a minimum post-evaluation fitness of 80%. Figure 8 shows the post-evaluation results for the controller which resulted in the highest post-evaluation fitness for each experiment.
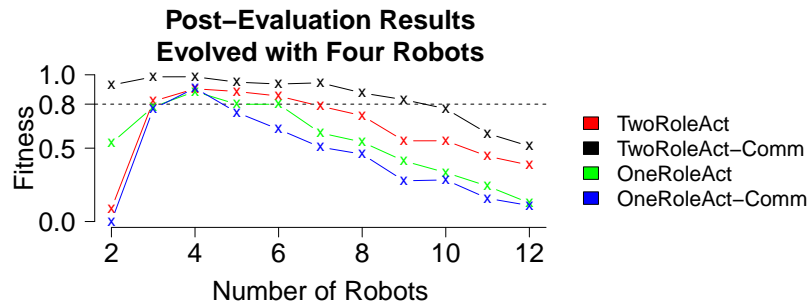

Figure 8: Post-Evaluation results for the three-role experiments.

For OneRoleAct and TwoRoleAct architectures we evolved controllers able to obtain 80% of post-evaluation fitness with four robots. However, only the TwoRoleAct controllers show scalability. With no signal differentiation reward, the TwoRoleAct controller scales to groups of three to six robots and when using the signal differentiation reward, TwoRoleAct-Comm, scales to groups of two to nine robots.

4.1.2   Behavior

The behavior of the highest post-evaluation fitness controllers, for ten trials, is shown in Figures 9 and 10. For OneRoleAct, decimal numbers represent the values detected by the role sensors. For TwoRoleAct, decimal numbers represent the values being emitted by the role actuators. The paths followed by the robots during the ten trials are represented by gray trails.

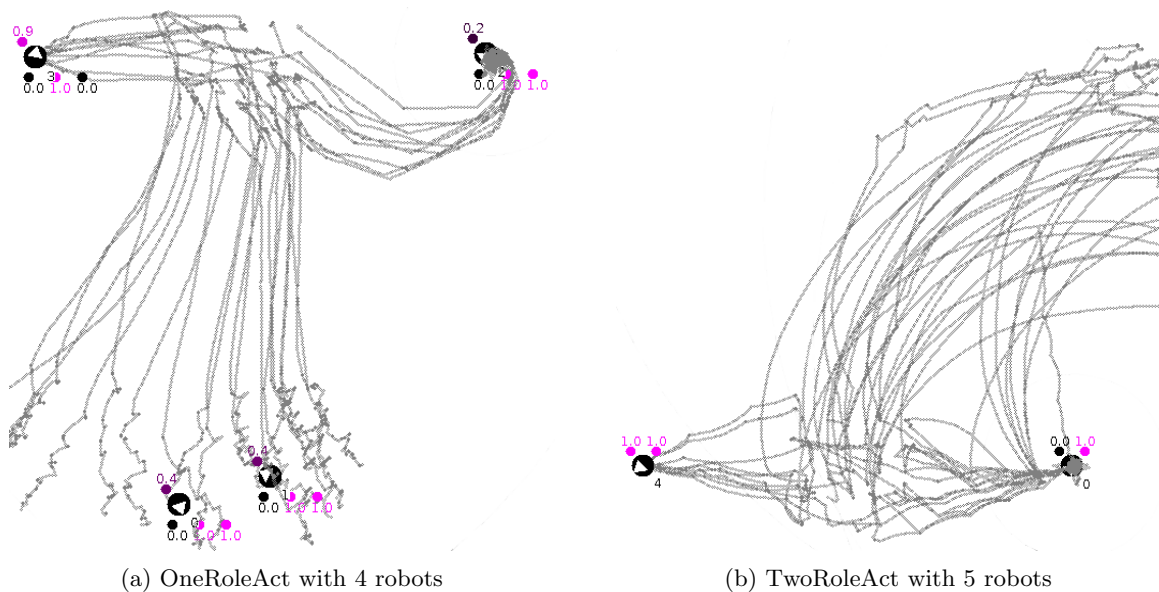(a) OneRoleAct with 4 robots (b) TwoRoleAct with 5 robots

Figure 9: Ten trials paths for OneRoleAct and TwoRoleAct with no signal differentiation reward.

For the OneRoleAct controller (Fig. 9 (a)), one robot emits a high value (≈ 1.0) and moves to the left light – at figure top left – while the other robots move South, until one of them emits a low value (≈ 0.2) and moves to the right light – at figure top right. The robots moving South emit ≈ 0.5 and stop at a distance where the lights are within the light sensor range. For the TwoRoleAct controller (Fig. 9 (b)), one robot emits high values (≈ 1) on both channels and moves to the left light while other robot emits low values (≈ 0.0) on both channels and moves to the right light – the two robots at figure bottom. The other robots emit a high value (≈ 1.0) on one channel and a low value (≈ 0.0) on the other channel and move North-East, to a distance where the lights are within the light sensor range – outside the figure. For OneRoleAct and TwoRoleAct, when the robots reach the positions at the lights, communication becomes unstable and we are unable to identify a pattern.

(a) OneRoleAct-Comm with 5 robots   (b) TwoRoleAct-Comm with 7 robots
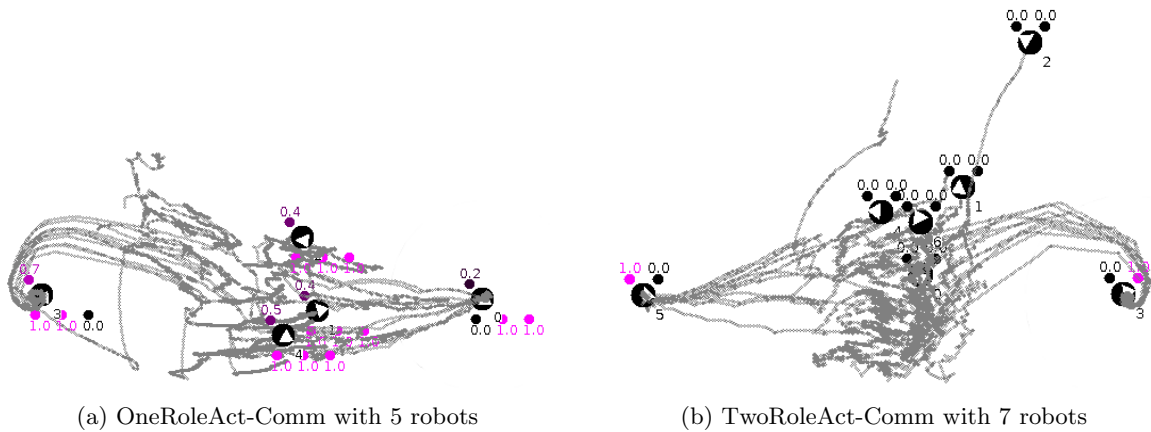
Figure 10: Ten trials paths for OneRoleAct and TwoRoleAct with signal differentiation reward.

For OneRoleAct-Comm and TwoRoleAct-Comm, a trial begins with a negotiation phase where robots allocate roles. For the OneRoleAct-Comm controller, one robot emits a value in $[0, \frac{1}{3}[$, another robot emits a value in $[\frac{2}{3}, 1]$ and all other robots emit a value in $[\frac{1}{3}, \frac{2}{3}[$. After the negotiation phase, the robots emitting in $[\frac{1}{3}, \frac{2}{3}[$ decrease motion, while the other two robots separate from the group and move to the lights. For the TwoRoleAct-Comm controller, the negotiation phase proceeds until one robot emits a high value (≈ 1.0) with one of the two role actuators, another robot emits a high value (≈ 1.0) with the other role actuator and all other robots emit low values (< 0.4) in both role actuators. After the negotiation phase, the robots emitting high values move to the lights while the other robots decrease motion, or exceptionally move about 1 m North.

Results show that robots learn to associate communication patterns with behaviors. For instance, in TwoRoleAct, emitting a high value on a channel might mean moving to the right light and in OneRoleAct, receiving data simultaneously in all channels means to avoid any light. Furthermore, the relationship between a communication pattern and a behavior is evolved differently for different runs, suggesting that evolution is finding different ways to use communication for role allocation.

5   Discussion and Future Work

We substitute the fixed-topology neuroevolution algorithm, previously used in [9], by NEAT, which is a neuroevolution algorithm that evolves both topology and weights, and we are able to evolve scalable strategies where previously had not been possible. Our results suggest that future research should avoid fixed-topology neuroevolution algorithms.

For the two-role task, we show how to evolve communication-based scalable role allocation strategies without rewarding signal differentiation. We introduce a novel fitness function that does not reward signal differentiation, and demonstrate that it is simple to co-evolve the necessary communicative and non-communicative behavioral aspects, contrary to previous findings [9]. Furthermore, our fitness function is more adequate to measure the quality of the evolved strategies, as it does not suffer from the limitations

identified in fitness functions used in earlier research [9]. We show that although communication is not strictly necessary to perform the task, it is a relevant factor to evolve scalable solutions.

Evolving communication is not trivial because evolution must produce both appropriate signals and corresponding reactions [1]. In our research, however, for the two-role task and for the three-role task, we evolve a communicative system without explicit selective pressure for communication use in the fitness function. Nevertheless, for the three-role task, the evolved communicative system shows limited scalability when the reward for signal differentiation is not present. When the signal differentiation reward is present, the evolved controllers show higher post-evaluation performance and scalability. Advances in the field of evolutionary robotics are achieved, amongst other ways, by designing systems able to perform increasingly complex tasks while minimizing the amount of a priori knowledge from the designer [4] [14]. We will research a fitness function that does not require rewarding signal differentiation, to avoid setting a priori a communicative strategy. Focus should be on defining fitness functions that accurately measure the quality of the evolved strategies and not how communication is used.

The presented results show that communication based role allocation might be useful as a role allocation strategy for two and three roles as well as for different numbers of robots. Although we show how two channels might be used to perform a three-role task, the relationship between the number of roles and the minimum number of channels is unclear. Our research suggests that one more channel is needed for every added role, which forces the designer to define a priori the number of roles and channels, thus limiting evolutionary exploration and the discovery of novel solutions. Therefore, we will research a higher decoupling between the number of roles and channels. Our aim is to build a cognitive architecture which allows a low number of channels and a high number of roles, possibly by combining the information on different channels.

Communication-based role allocation is a research path worth pursuing because it might offer scalability and robustness for cooperative multi robot systems. Our goal is to find a generalizable evolutionary setup to evolve scalable and robust communication-based role allocation. Therefore, we will research an evolutionary setup able to evolve strategies for three-role tasks without explicit selective pressure for communication use in the fitness function. We aim to identify the conditions for the emergence of communication-based role allocation strategies for increasingly larger numbers of roles and robots.

Acknowledgements

References

[1] Christos Ampatzis, Elio Tuci, Vito Trianni, and Marco Dorigo. Evolution of signaling in a multi-robot system: Categorization and communication. *Adaptive Behavior*, 16(1):5–26, 2008.

[2] G. Baldassarre, S. Nolfi, and D. Parisi. Evolving mobile robots able to display collective behaviors. *Artificial Life*, 9(3):255–267, 2003.

[3] Mohammad Divband Soorati and Heiko Hamann. The effect of fitness function design on performance in evolutionary robotics: The influence of a priori knowledge. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, GECCO'15, pages 153–160, New York, NY, 2015. ACM.

[4] Stephane Doncieux and Jean-Baptiste Mouret. Beyond black-box optimization: a review of selective pressures for evolutionary robotics. *Evolutionary Intelligence*, 7(2):71–93, 2014.

[5] Miguel Duarte, Fernando Silva, Tiago Rodrigues, Sancho Moura Oliveira, and Anders Lyhne Christensen. JBotEvolver: A versatile simulation platform for evolutionary robotics. In *14th International Conference on the Synthesis and Simulation of Living Systems–ALIFE*, pages 1–8, New York, NY, 2014. MIT Press, Cambridge, MA.

[6] D Floreano and J Urzelai. Evolutionary robots with on-line self-organization and behavioral fitness. *Neural Networks*, 13(4â5):431–443, 2000.

[7] Onofrio Gigliotta. Task allocation in evolved communicating homogeneous robots: The importance of being different. In *Trends in Practical Applications of Scalable Multi-Agent Systems, the PAAMS Collection*, pages 181–190, Cham, Switzerland, 2016. Springer International Publishing.

[8] Onofrio Gigliotta, Marco Mirolli, and Stefano Nolfi. Who is the leader? dynamic role allocation through communication in a population of homogeneous robots. *Artificial Life and Evolutionary Computation*, pages 167–177, 2010.

[9] Onofrio Gigliotta, Marco Mirolli, and Stefano Nolfi. Communication based dynamic role allocation in a group of homogeneous robots. *Natural Computing*, 13(3):391–402, September 2014.

[10] L. König. *Complex Behavior in Evolutionary Robotics*. Walter de Gruyter GmbH & Co KG, 2015.

[11] Gustavo Martins, Paulo Urbano, and Anders Lyhne Christensen. Using communication for the evolution of scalable role allocation in collective robotics. In Guillermo R. Simari, Eduardo Fermé, Flabio Gutiérrez Segura, and José Antonio Rodríguez Melquiades, editors, *Advances in Artificial Intelligence - IBERAMIA 2018*, pages 326–337, Cham, 2018. Springer International Publishing.

[12] Francesco Mondada, Michael Bonani, Xavier Raemy, James Pugh, Christopher Cianci, Adam Klaptocz, Stephane Magnenat, Jean-Christophe Zufferey, Dario Floreano, and Alcherio Martinoli. The e-puck, a robot designed for education in engineering. volume 1, pages 59–65, Portugal, 2009. IPCB:Instituto PolitÃ©cnico de Castelo Branco.

[13] NEAT4J. Neat4j Java framework, 2006.

[14] Andrew L. Nelson, Gregory J. Barlow, and Lefteris Doitsidis. Fitness functions in evolutionary robotics: A survey and analysis. *Robotics and Autonomous Systems*, 57(4):345–370, 2009.

[15] Stefano Nolfi. Evolutionary robotics: Exploiting the full power of self-organization. *Connection Science*, 10:167–183, 1998.

[16] Stefano Nolfi and Dario Floreano. *Evolutionary Robotics: The Biology, Intelligence, and Technology*. MIT Press, Cambridge, MA, USA, 2000.

[17] Matt Quinn, Lincoln Smith, Giles Mayley, and Phil Husbands. Evolving controllers for a homogeneous system of physical robots: structured cooperation with minimal sensors. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 361(1811):2321–2343, 2003.

[18] Kenneth O. Stanley and Risto Miikkulainen. Efficient reinforcement learning through evolving neural network topologies. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2002)*, page 9, San Francisco, CA, 2002. Morgan Kaufmann.

[19] Kenneth O. Stanley and Risto Miikkulainen. Evolving neural networks through augmenting topologies. *Evolutionary Computation*, 10(2):99–127, 2002.

[20] Elio Tuci, Boris Mitavskiy, and Gianpiero Francesca. On the evolution of self-organised role-allocation and role-switching behaviour in swarm robotics: a case study. *Procdings of the European Conference on Artificial Life (ECAL 2013)*, pages 379–386, 2013.