Autonomous Learning of collaboration among robots

Paolo Arena, Luca Patané and Alessandra Vitanza

***Abstract*— The aim of this paper is to study the emergence of coordinated activities, and the investigation of collaboration between individuals in a small group of robots. The idea is to impose very simple global rules and to give a primary role to the environment mediation. In the paper the specialization strategy, already introduced in a previous work is extended, to autonomously solve a task assignment problem among agents in an initially homogeneous swarm. In particular, a given sequence of tasks is assigned to the group and each robot has to autonomously specialise in solving sub-sequences, resulting in a labor division which improves the performance of the team. Behavioral improvement is guided by a global reward function. Results, obtained in a dynamic simulation environment, show that performances depend by environmental conditions and starting positions of the singular agents: environment and the other robots play clearly a fundamental role in mediating the swarm capabilities.**

I. Introduction

In literature artificial swarm systems are seen as autonomous aggregation of homogeneous or heterogeneous agents with distributed and self-organizing capabilities, sometimes with local communication mechanism, often without direct but environmentally mediated communication. Thanks to their characteristics these systems result robust for adapting to environmental changes [1], [2]. Adaptation is the mechanism produced by the modulation of individual (e.g. insect) behaviors and it is very important for recovery and fault tolerance aspects. A typical example of adaptation is the capability of ants to modify their paths in order to adapt to environmental changes when solving a given task.

Among peculiarities detectable in biological swarm systems *Coordination* and *Collaboration* are the two crucial characteristics investigated in this paper. The former is seen as the spatio-temporal distribution among the individuals, of their activities and of the tasks required to resolve complex problems, whereas the latter is identified as the capability to perform several activities simultaneously by groups of specialized individuals [3]. It is due to a morphological and natural differentiation influenced by the age of the individuals that permits the specialization of different behaviours. An example of coordination is the exploitation of the pheromone trail used by ants [4]. Another example of coordination is the organization of the displacement in bee and locust swarms, where different members interact together to realize a temporal (synchronization) and a spatial (orientation) arrangement toward a specific goal. As an example of Collaboration, bees are specialized in different activities obtaining an unique emergent behaviour: for brood disposing inside the nest and foraging, some agents go out to search food, whereas other members stay and work at the nest [3].

Relating to the evolution strategy, conceptually, it is possible to choose between two different approaches: Phylogenesis and Ontogenesis. Whereas the classical swarm intelligence algorithms are based on the former concept, involving small and incrementally gradual changes during evolution, the strategy adopted in this work is rather based on the latter: changes are applied to the same genetic structure within the agent and learning improves the agent capabilities within its limited lifetime. This is much more efficient when looking for a strategy where adaptation is required from the scratch in a group of robot initially endowed with no knowledge about how to accomplish a given task.

As this paper will experimentally demonstrate, the development within each agent, mediated by the environment, even without the interaction with the other agents, involves a move from simpler, un-organised behaviors to more complex ones, leading to the specialization of each agent to respond efficiently to its own experiences contributing, at the same time, to receive the global reward.

Our approach is based on a reinforcement learning, and a global rewarded signal is activated when the overall mission is reached. An important aspect that our approach can efficiently deal with and that will be further explored in the near future, is the concept of *Diversity*, seen as the possibility to have situations where autonomous agents have the same mechanical structure, but differ in performing behaviors. This characteristic permits to achieve an evolutionary mechanism in a behaviorally heterogeneous group from an initially homogeneous population. Moreover, whereas diversity deals with difference among agents no matter the performance improvements, specialization is seen as an adaptation of inborn structures in order to fit a specific role which enhances performance. The combination of Diversity and Specialization permits the system to become diverse ensuring better performance [5]. This is also in line with biological inspiration where the specialization is also influenced by the structure of the individuals and it emerges from a behavioral and morphological differentiation.

II. The proposed approach

The proposed approach follows previous experiences where a new learning method was drawn [6] to allow the agents to detect and reach all the targets in the environment. The strategy is here extended towards the formation of swarming capabilities, with particular attention to the emergence of collaboration. For this reason, in our scenario we considered an autonomous group of agents, which start

Paolo Arena, Luca Patané and Alessandra Vitanza are with the Dipartimento di Ingegneria Elettrica, Elettronica e Informatica, Universitá degli Studi di Catania, Italy (email: {parena, lpatane, avitanza}@dieei.unict.it).

with the same cognitive knowledge and concurrently act in the same environment where a series of differently colored targets are present on the floor. These are cyclically activated, once at a time. When a robot reaches an active target, this one disappears and the following target is activated becoming visible in the scene. A global reward signal is forecasted to all the agents whenever the last target is reached by one robot: only this event induces learning in robots. During the learning phase each robot, due to its random motion in the arena, visits a different sequence of targets, so, the activation of the reward signal biases its behavior according to its own experiences. In case of reward, in fact, each robot increases its willingness to reach the visited color targets and decreases its interest in the others color targets.

The final expected situation is to obtain decoupling of color sensitivities for each robot, to perform collaborative sub-tasks and optimize efforts. Thanks to the flexibility of our neural architecture, agents learn to collaborate performing specialization through an emergent labor division to reach a common intent. Starting from a homogeneous indirect communicating team, in which each robot has the same mechanical body and neural control structure, robots are induced to collaborate to achieve an overall global task guided by the reward-based learning mechanism.

A. Neural Model description

The design of the learning mechanism, above mentioned, is based on a previous formalization [6], which uses a correlation-based learning approach in bio-inspired spiking networks [7] [8]. The neuron model chosen for implementation is the Class I Izichevich neuron, considered as a good trade-off between computational load and biological plausibility. The class I was adopted because the spiking rate is proportional to the amplitude of the stimulus [9]: this configuration is suitable for sensing neurons. Regarding the model of synaptic learning, instead, the Hebbian learning method (STDP - Spike Timing Dependent Plasticity) was used to create correlations between Unconditioned Stimuli (US) and Conditioned Stimuli (CS). In order to permit the new feature, a threshold plasticity was added to the STDP, where the modulation of the threshold was obtained through the adaptation of the presynaptic bias current present in the model of the spiking neuron.

The proposed network is composed of two main blocks: the subnetwork dedicated to obstacle avoidance (Fig.1(a)) and the subnetwork related to visual target recognition (Fig.1(b)).

The inter-correlation between both sub-networks permits to control the robot movements using a specific transduction function to transform the output of motor-neurons into the input to the motors of the simulated robot. The object approaching block permits the robot to identify specific targets in order to reach them. It is composed of a certain number of neurons organized in two levels and interconnected to analyze the image obtained by a simulated camera and identify the target position in the field of view.

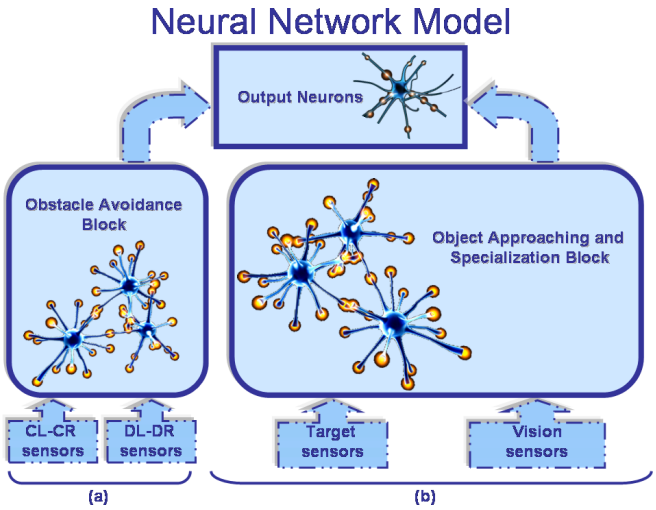


Fig. 1. A block diagram of the neural network used to control each robot: **(a)** Obstacle avoidance subnet - inputs: information from contact CL-CR sensors (US) and distance DL-DR sensors (CS). To guarantee rapid obstacle avoidance actions, this subnet has major priority and is not subject to learning. **(b)** Object approaching subnet - inputs: data from target-sensors (US) and vision sensors (CS), used to find the position of targets. This subnet has to be replicated as many times as the number of different features to recognize. Finally, the upper block contains motor neurons outputs used to directly control the wheels of the robots.

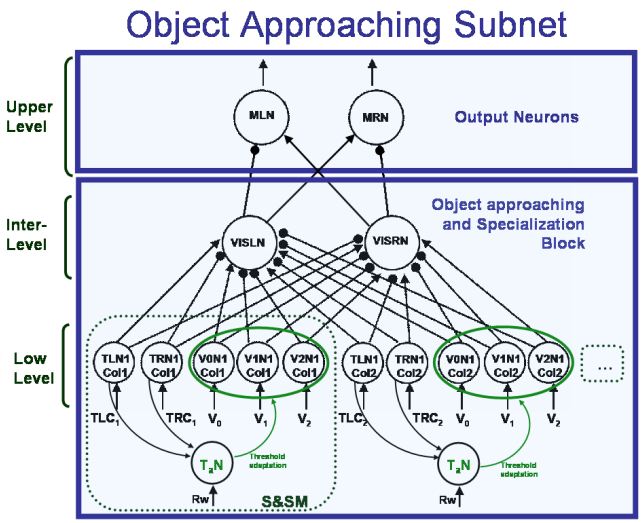


Fig. 2. Details about the Object approaching subnet: the structure of this part is composed of a two-layers neural network. TLC_x/TRC_x ($x = 1, 2..n$) are the inputs coming from target sensors and used as unconditioned stimuli (US). $V_0/V_1/V_2$ are inputs coming from vision sensors used to guide the approach of this target. This structure must be replicated for all targets. All these neurons act on the motor neurons (*MLN/MRN*) through the inter-neurons *VISLN/VISRN*. Output neurons permit to directly control the wheels on the left (*MLN*) and right (*MRN*) side of the robot. T_aN represents the Threshold adaptation neuron. It is responsible of the activation of Specialization learning.

Referring to Fig. 2, the first layer is constituted by sensory neurons that are connected to motor-neurons through a layer of inter-neurons. At the beginning these stimuli are able to trigger a response for all targets present in the scene. This is due to pre-learned paths from the conditioned sensory layer

to the motor layer. The network was already trained, using the STDP rule, only to outline the effect of specialization. In [6] it was demonstrated that STDP learning and role specialization learning can coexist. The final aim of the experiment is to obtain different robots specialized to reach different sub-groups of the target set: the division of duties is needed in order to complete the task in a better way.

To comprehend the mechanism used to realize the threshold adaptation, the neuron model used is reported below[9]:

$$\begin{aligned} \dot{v} &= 0.04v^2 + 5v + 140 - u + I \\ \dot{u} &= a(bv - u) \end{aligned} \tag{1}$$

with the spike-resetting

$$\texttt{if } v \geq 0.03, \texttt{ then} \begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases} \tag{2}$$

v, u represent, respectively, the neuron membrane potential and the recovery variable. I represents the pre-synaptic input current, finally a, b, c and d are system parameters and in particular, $a = 0.02$, $b = -0.1$, $c = -55$, $d = 6$. The time unit is ms.

Starting from this model, the input current I (in eq. 1) is split in two contributes: I_i, seen as input from sensorial and synaptic inputs, and a term modeled as a voltage-dependent current. It is seen as a bias subject to the adaptation effects and its fluctuations depend on the reward-signals received. In particular, defining g_A as an activation-conductance, this term can be expressed as $g_A V_{thresh}$ [10]. Equation (1) becomes:

$$\begin{aligned} \dot{v} &= 0.04v^2 + 5v + 140 - u - g_A V_{thresh} + I_i \\ \dot{u} &= a(bv - u) \end{aligned} \tag{3}$$

It has to be noticed, as outlined below, that positive variations in V_{thresh} produce a neuron facilitation, whereas negative ones cause hyperpolarization. In our model we use $g_A = 1$. For a given robot, the Object approaching subnet consists of an input layer organised into *sensing and specialization modules* (S&SM), each one constituted by neurons dealing with unconditioned (target) sensors and with conditioned (vision) ones (see Fig.2). Within the sensing block associated to each color target the neuron dealing with conditioned stimuli undergo threshold adaptation thanks to the action of the *threshold adaptation Neuron (T_aN)*, which adds a contribution in the threshold V_{thresh} for the vision neurons for all the sensing modules within the object approaching subnet, only if the unconditioned neurons within the S&SM and the Reward (Rw) signal are both active, following the formula below:

$$\Delta V_h = \begin{cases} \Delta V_h & Rw > 0 \;\; and \;\; V_{TN} > V_{TNthresh} \\ 0 & otherwise \end{cases} \tag{4}$$

The corresponding threshold modulation for the vision neurons is:

$V_{thresh} = V_{thresh} + \Delta V_h$

This adaptation acts twofold: $\Delta V_h = \Delta V_D$ as a depolarization effect on the vision neurons within the same S&SM, and $\Delta V_h = \Delta V_H$ as a hyperpolarization for the vision neurons within the other S&SMs (see Fig. 2 for details). In our simulations we adopted $\Delta V_D = 1.8$; $\Delta V_H = -0.6$. This adaptation takes place only after a reward signal is generated: for each robot, vision neurons for all the active and visited targets are depolarized, whereas all the other ones within the other S&SMs are hyperpolarized.

Since the interrelations among members of the swarm are mediated by the reward function without direct-communication among the robots, the reward signal acts as an external global input for all agents. In this scenario, a single agent can see only local information and no global situation.

In our scenario, as already outlined, the robots are already trained to avoid obstacles and reach targets [8], [6], in order to emphasize the emergence of cooperation abilities. The threshold adaptation strategy was applied to promote the evolution of emerging capabilities among robots, in line with ecological results where colonies take a profit by the possibility to distribute information [11].

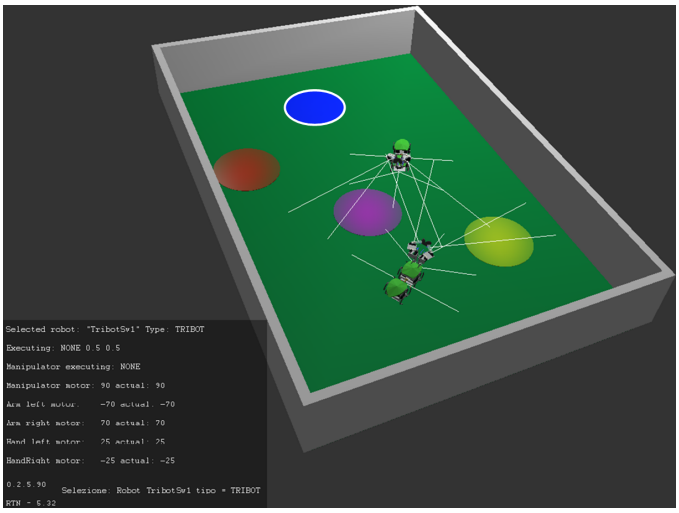
III. Simulation results

As in our previous works, simulations to validate the proposed approach were performed using the SPARKRS4CS architecture. It is a software/hardware framework realized to develop cognitive systems in union with an ad hoc Dynamic Robotic Simulator, more details in [12]. The simulated environment used for robot experiments is an arena (3m x 2m) with a number of different target areas alternately enabled on the floor. The model of the robots implemented is the simulated version of a bio-inspired hybrid mini-robot, called TriBot I [13], realized within the European projects SPARK I-II [14], [15]. The particular design of the TriBot permits to overcome irregular terrains and climb high obstacles [16]. For this robot, the feasibility of an on board STDP learning scheme was also demonstrated in [8].

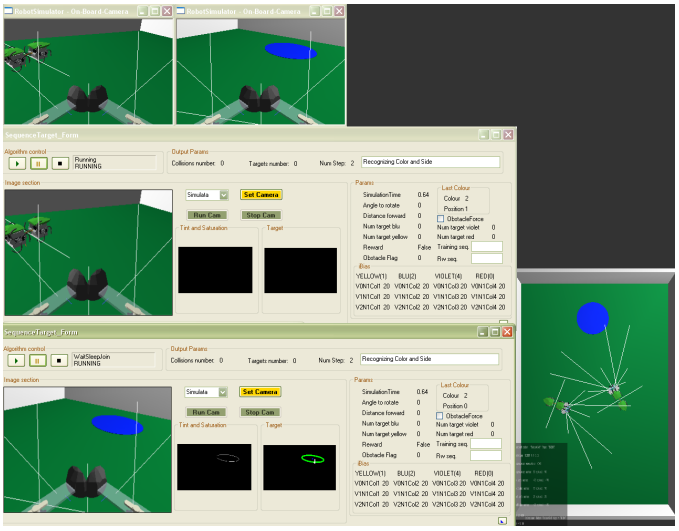
Fig. 3(a) shows an overview of the overall simulation environment used to implement the experiments, whereas Fig. 3(b) shows the control algorithms' interface.

Giving more details about the simulation settings and strategy, at the beginning of the test only a target (the blue-target in Fig.3) is enabled and visible on the floor. When a robot reaches it, the following target is activated and the previous one disappears from the scene. Whenever all targets are sequentially enabled and reached by robots, the reward signal is activated to induce learning in all individuals. After that, the situation is restored to original configuration and the cycle repeats. In particular, in these preliminary simulations we used combinations of two robots and four targets in different arrangements. The target-chain is composed of (starting from the first to the last enabled): blue, red, yellow and violet. Upon reaching the latter, the reward signal is broadcasted to all the robots. At implementation level a hearing source or a flashing light are just some examples signals stimulating a reward or a punishment.

These simple rules allow the robots to independently create sequences of visited targets, although no communication is



(a)



(b)

Fig. 3. **(a)** An image of the arena designed with the Dynamic Robotic Simulator. Two TriBot robots move in the environment with four targets-chain. The blue target (with white edge) is the first in the chain and it is enabled at the beginning of the experiment (as shown in (b)). The other targets (faded in figure) are going to be enabled whenever a robot reaches the preceding target in the chain. **(b)** The GUI software architecture (SPARKRS4CS). The graphical interface related to the two robots and the on-board cameras are shown on top of the figure.

furnished.

In these preliminary experiments the results related to the capabilities of threshold adaptation to induce learning of sub-sequences of targets are evaluated. Initially each robot is sensitive to all targets, characterized by different colors (Col1 = yellow and Col2 = blue Col3= red Col4 = violet). At the end of the specialization learning, robots are interested only in a sub-set of targets.

Starting from the same value $g_a V_{thresh} = 20$ all bias

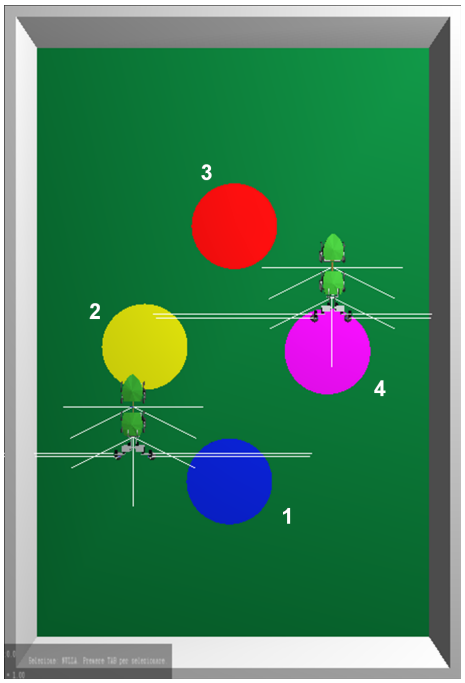


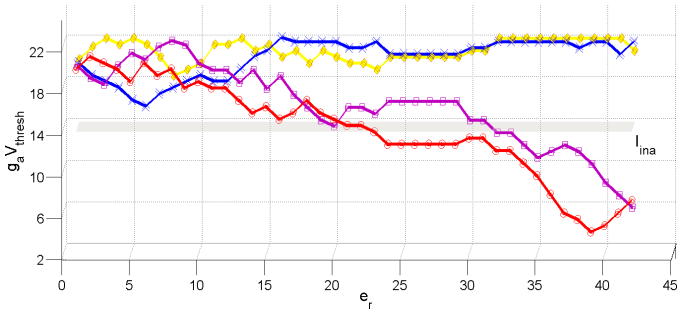
Fig. 4. Screenshot of the experimental setup for Simulation 1. Numbers indicate the activation order of the targets in the chain.

currents related to the different neurons can be modulated, considering the following saturation values: $0 \leq g_a V_{thresh} \leq 22$. The threshold adaptation is considered complete when the lower bound, called inactivation current $g_A V_{thresh} = I_{ina} = 14$ is reached. In these conditions, even if an input current is present, the total current value cannot overcome the threshold and the neuron is no longer sensitive to external stimuli.

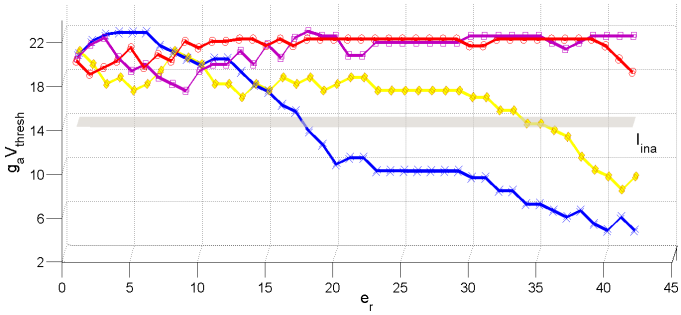
Fig. 5(a) and Fig. 5(b) show the dynamic fluctuation of the bias current $g_a V_{thresh}$ due to this mechanism. In the time window, after about 20 reward activation events (e_r in Fig. 5), Robot 1 learns to focalize its attention on blue and yellow targets, whereas the emerging behaviour of Robot 2 is a specialization in red and violet targets. Starting position of the robots and the arrangement of targets for this simulation are shown in Fig.4.

In particular, Fig. 5 shows the neuron current of one of the Vision neurons devoted to a particular color target. It is clear that at the beginning of the experiment all current values are modulated but currents are maintained at high values since all robots are pre-trained to recognize and approach all targets, whereas after about 10 reward activations, robots start to show different specialization behaviours. Robot 1 proceeds reaching more frequently the yellow and blue target, whereas the thresholds of the neurons related to the other color targets are decreased. During learning, for these last neurons it becomes more and more difficult to exceed the threshold value. A similar situation involves in robot 2, but regarding the other targets.

A number of different environments were simulated, involving different arrangements of the arena, displacement



(a)



(b)

Fig. 5. **(a)** - Robot 1 dynamic fluctuations of the bias current. This robot is going to specialize in blue and yellow targets; so as shown in figure after approximately 20 reward activations, the bias current of red and violet targets go below lower bound value. **(b)** - Robot 2 dynamic fluctuations of the bias current. This robot becomes sensitive to red and violet targets, so in this case the current related to the blue target needs 17 rewards to overcome the lower bound value (I_{ina}), whereas bias current of the yellow target comes down bound value after 35 rewards.

of the targets and initial position of the robots. As another example, results of a different simulation are reported in Fig. 7, to demostrate the validity of the approach and to detect similarities and differences through comparisons. This simulation differs in starting position of the robots and in the arrangement of targets as shown in Fig.6.

Fig. 7 shows the neuron current for the same Vision neuron as in Fig. 5. At the beginning all currents keep high values, but in this case just after about 5 reward activations, robots start to show specialization effects. In this simulation, the robot 1 proceeds specializing in blue and violet target, whereas the robot 2 in yellow and red ones. Similar to the previous results is the trend of bias current g_aV_{thresh} but this experiment needs about 23 reward activation events to complete the theshold adaptation.

Fig. 8(a) and Fig. 8(b) show the trend of the number of targets found by the two robots in relation to Simulation 1. It can be noticed the difference between the color targets in which each robot is going to specialize and the remaining targets that the robot will incrementally ignore.

Referring again to Simulation 1, Fig. 9(a) reports the first event of activation of the reward function. Robot 1 trajectory

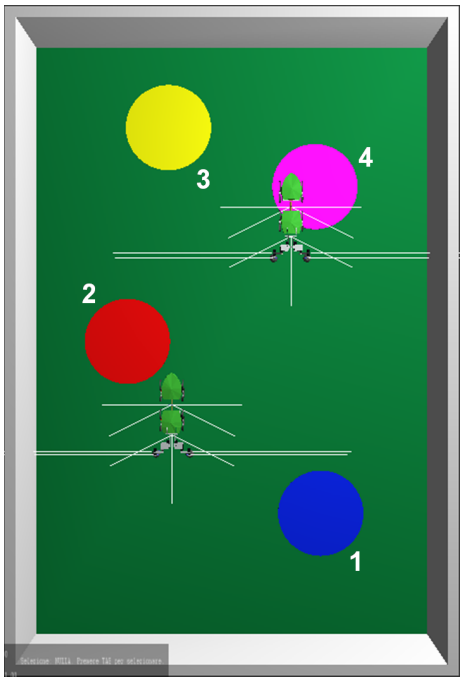
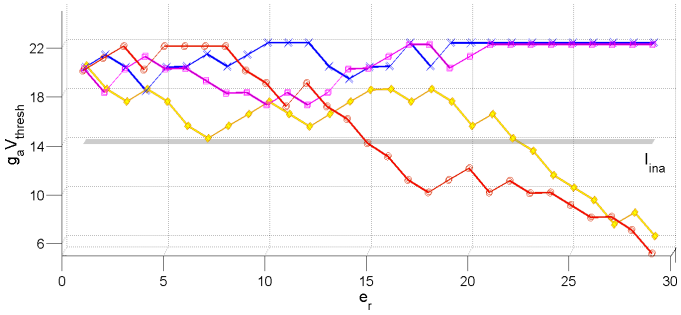


Fig. 6. Environmental setup for Simulation 2. In this arena the targets are more distributed. The order of activation for the targets is shown.

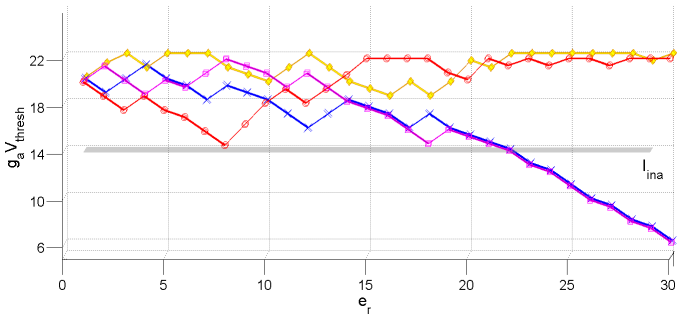
is reported in blue and outlined with the blue bullet with the robot identification number, whereas robot 2 path is depicted in green. As it can be noticed, both the robots are attracted by all the targets. Environmental conditions and the initial robot position have the important role of the initial biasing of the learning phase. In fact as it can be seen from the figure, according to the sequence, the blue target (first activated in the chain) attracts both the robots, but the nearest one (robot 2) arrives before the other (at robot step number 2). This event leads to the activation of the second target (yellow one). This is perceived by both robots, but robot 1 arrives slightly before the other. Robot 2 has then to avoid robot 1 soon after step 13. In the mean time the red target is activated and, once again, robot 1 reaches it, activating the last target, which is reached by robot 2. At this time the reward signal is broad casted to all the robots, and the sequence just visited by each robot is reinforced. The activation of the reward function at the event $e_r = 40$ is reported in Fig. 9(b). Here the task division is clearly visible: robot 1 is specialised in reaching violet and red targets, whereas the other robot prefers the other targets. It is to be noticed that robot 2 passes through the yellow target when this is not activated: the robot simply follows the path to reach the red target.

IV. Remarks and Future Works

The possibility to use an adaptation mechanism, which is biased toward exploiting the capabilities of each individual to induce specialization in an group of robots is an interesting approach to permit the emergence of collective behaviours and division of labour. The key remark to be underlined is that in this work no particular capabilities are ascribed to

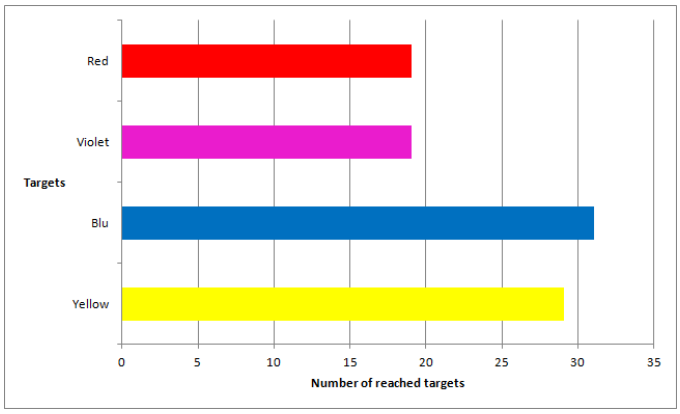


(a)

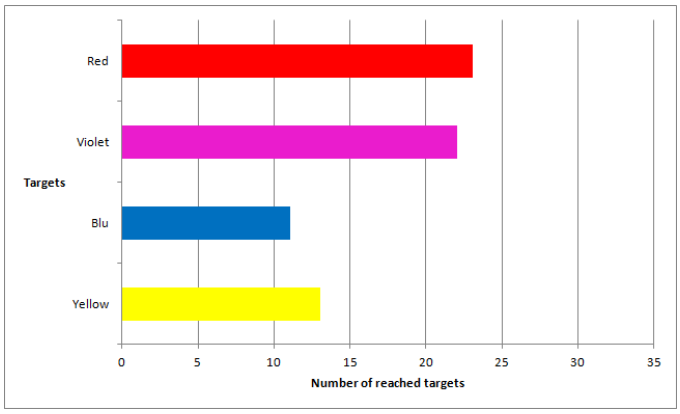


(b)

Fig. 7. Further experiment; **(a)** - Robot 1 dynamic fluctuations of the bias current. This robot is going to specialize in blue and violet targets; after approximately 15 reward activations, the bias current of red targets go below lower bound value, whereas the bias current of yellow target needs about 23 reward activation events (e_r) to learn completely. **(b)** - Robot 2 dynamic fluctuations of the bias current. This robot remains reactive to red and yellow targets. In this case the current of the blue and violet targets after about 17 rewards overcomes the lower bound value (I_{ina}) and the specialization can be considered completed.



(a)



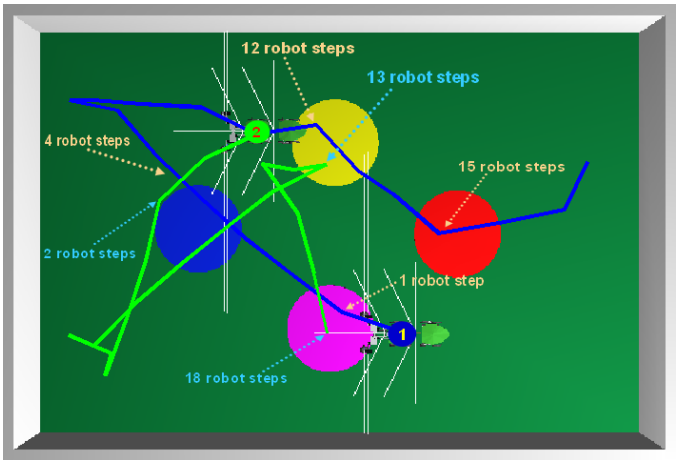
(b)

Fig. 8. Cumulative number of visited targets: **(a)** related to the robot 1, which specializes in yellow and blue targets, **(b)** related to the robot 2 which specializes in red and violet targets.

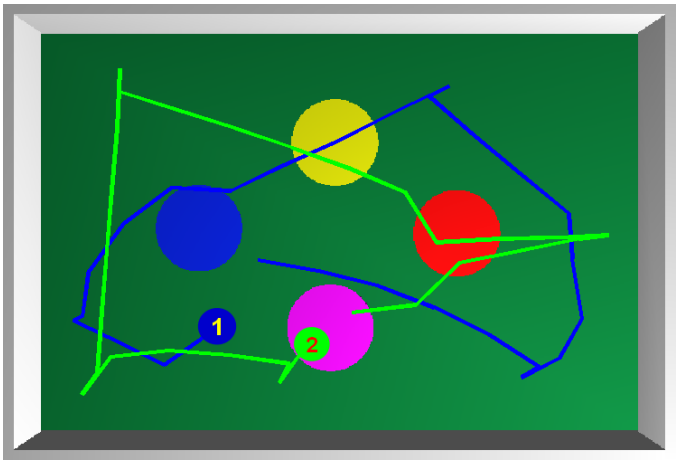
each agent within the group. Even in this case, a suitable task division among the agents is obtained, exploiting the mediation of the environment through the action of the reward function.

The simulation results obtained demonstrate that the presence of a global reward induces diversity in a team of homogeneous robots as also discussed in general approaches to swarm intelligence [17], [18]. On the other hand, it is not quite rare to find such a strategy in living swarms. Different studies [19], [20] show how various species of social insects use an interesting communication system based on multimodal signals to exchange information through the colony. They use pheromones, visual, acoustic and tactile signals that reach a large amount of the colony population.

Moreover, as it can be seen from Fig.9(b), the simple addition of a visual scanning to detect the active target before moving would really decrease the number of robot steps needed for the reward activation. Starting from these preliminary results, it is possible to underline a lot of future scenarios where, for example, the robots could be endowed with different and much more enhanced individual capabilities, or groups of robots with the same reward function but different robotic structures or instead different rewards acting on the same robotic structure. For instance, it could possible to use the same cognitive architecture inside different robotic platforms like wheeled and legged robots, robotic arms, hybrid vehicles. Thus, a heterogeneous group of robots can be designed, endowed with the same cognitive architecture but with different preferences and priorities inside. Consequently, for example, if a wheeled robot without manipulator can prefer, at first, to reach a manipulable target, during the learning phase, it will acquire knowledge about its inability to satisfy its priority demand, so it can change and adapt its objectives in order to specialize itself in a more useful role within the swarm. This is on the line of biological inspiration, where specialization is influenced by the structure of the individuals and it emerges from a behavioral and morphological differentiation. Furthermore, if different kinds of rewards are introduced, individually for each robot it is possible to elicit different behavioral



(a)



(b)

Fig. 9. Simulation 1: **(a)** Trajectories performed by the robots before the first activation of the reward signal. For the sake of clarity the robot steps number rs when targets are enabled T_{en} and disabled T_{dis} are reported. blue target: $T_{en} = 0\ rs$, $T_{dis} = 2\ rs$, yellow target: $T_{en} = 2\ rs$, $T_{dis} = 12\ rs$, red target: $T_{en} = 12\ rs$, $T_{dis} = 15\ rs$, violet target: $T_{en} = 15\ rs$, $T_{dis} = 18\ rs$. **(b)** Trajectories performed by the robots during the last activation of the reward signal.

responses according to each robot structure, obtaining a natural and morphologically-based division of labor.

The possibility to learn focalizing attention in a sub-set of alternatives allowed, suggests future scenarios when sequences of actions are associated to the learned sub-set. For instance, if a communication language is established, each robot can learn if a reward is correlated with another robot action. In this way, a correlation-based association can be realized to converge to an optimized sequence of actions in the swarm.

Instead, in this scenario the final arrangement is affected strongly by environmental conditions and events, such as obstacle avoidances movements or exploration roundtrip.

V. Conclusion

In the paper an approach for the emergence of cooperation in a swarm of robots was refined; the formulation of a further control strategy based on a Role Specialization learning method has been introduced. A description of the proposed neural structure and the corresponding learning mechanisms is given. Moreover preliminary simulation results are achieved to justify and evaluate the relevance of the approach. Finally, some examples of emergent strategies have been outlined to identify possible future interesting investigations.

Acknowledgement

This work was supported by EU Project EMICAB, Grant N 270182.

References

[1] E. Bonabeau, M. Dorigo, and G. Theraulaz, *Swarm Intelligence: From Natural to Artificial Systems*. New York: Oxford Univ. Press, 1999.

[2] A. Farinelli, R. Farinelli, L. Iocchi, and D. Nardi, "Multi-robot systems: A classification focused on coordination," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, pp. 2015–2028, 2004.

[3] A. J. Ijspeert, A. Martinoli, A. Billard, and L. M. Gambardella, "Collaboration through the exploitation of local interactions in autonomous collective robotics: The stick pulling experiment," *Autonomous Robots*, vol. 11, no. 2, pp. 149–171, Sep 2001.

[4] C. Kube and E. Bonabeau, "Cooperative transport by ants and robots," *Robotics and Autonomous Systems*, vol. 30, no. 1–2, pp. 85–101, Jan 2000. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0921889099000664

[5] L. Li, A. Martinoli, and Y. S. Abu-mostafa, "Diversity and specialization in collaborative swarm systems," 2003.

[6] P. Arena, L. Patané, and A. Vitanza, "Spiking networks and the emergence of cooperation through specialization," *submitted in: Transactions on Systems, Man, and Cybernetics–Part B: Cybernetics*, 2011.

[7] P. Arena, L. Fortuna, M. Frasca, and L. Patané, "Learning anticipation via spiking networks: Application to navigation control," *IEEE Transactions on Neural Networks*, vol. 20, pp. 202–216, February 2009.

[8] P. Arena, S. de Fiore, L. Patané, M. Pollino, and C. Ventura, "STDP-based behavior learning on the TriBot robot," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 7365, may 2009, p. 736506.

[9] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Transactions on Neural Networks*, vol. 14, no. 6, pp. 1569–1572, Nov 2003.

[10] J. Benda, L. Maler, and A. Longtin, "Linear Versus Nonlinear Signal Transmission in Neuron Models With Adaptation Currents or Dynamic Thresholds," *Journal of Neurophysiology*, vol. 104, no. 5, pp. 2806–2820, Nov 2010. [Online]. Available: http://jn.physiology.org/cgi/doi/10.1152/jn.00240.2010

[11] A. Dornhaus and L. Chittka, "Information flow and regulation of foraging activity in bumble bees (bombus spp.)," *Apidologie*, vol. 35, no. 2, pp. 183–192, Mar 2004. [Online]. Available: http://www.edpsciences.org/10.1051/apido:2004002

[12] P. Arena, M. Cosentino, L. Patané, and A. Vitanza, "SPARKRS4CS: a software/hardware framework for cognitive architectures (invited paper)," in *in: Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 8068, Kos, Greece, 2011, pp. 8068A–18.

[13] P. Arena, L. Patané, M. Pollino, and C. Ventura, "Tribot: a new prototype of bio-inspired hy-brid robot," St.Louis,Missouri,USA, 2009.

[14] (2004–2007) SPARK: Spatial-temporal Patterns for Action-oriented perception in Roving robots. [Online]. Available: http://www.spark.diees.unict.it

[15] (2008–2011) SPARK II: Spatial-temporal Patterns for Action-oriented perception in Roving robots II: an insect brain computational model. [Online]. Available: http://www.spark2.diees.unict.it

[16] P. Arena, L. Fortuna, M. Frasca, L. Patané, and M. Pavone, "Implementation and experimental validation of an autonomous hexapod robot," in *in Proceedings of the IEEE International Symposium on Circuits and Systems*, Kos, Greece, 2006, pp. 401–406.

[17] T. Balch, "Communication, diversity and learning: Cornerstones of swarm behavior," in *Swarm Robotics*, ser. Lecture Notes in Computer Science, E. Sahin and W. Spears, Eds. Springer Berlin / Heidelberg, 2005, vol. 3342, pp. 21–30.

[18] L. Bayindir and E. Sahin, "A review of studies in swarm robotics," *Turkish Journal of Electrical Engineering*, vol. 15, no. 2, 2007.

[19] J. Billen, "Signal variety and communication in social insects," in *in: Proceedings of the Netherlands Entomological Society Meetings*, vol. 17, 2006, pp. 9–25.

[20] R. Hickling and R. Brown, "Analysis of acoustic communication by ants." *J Acoust Soc Am*, vol. 108, no. 4, pp. 1920–9, 2000.