

## **On-line distributed learning algorithm in swarm of cooperative robots**

Université Pierre et Marie Curie, Paris, France

Advisors: Nicolas Bredeche, Nicolas Maudet

Contact: (firstname.name)@upmc.fr

Swarm robotics designates large groups of robots with limited communication and capability, that coordinates on a local basis to achieve tasks that can be evaluated on a larger scale. Typical examples are: collective construction (Werfel et al. 2014), self-assembly into specific patterns (Rubenstein et al., 2014), collective transport and exploration (Bayindir, Sahin, 2007).

The main challenge in swarm robotics is: how to program the behaviour of each of the robots, given the objective is defined at the level of the population (e.g. the total number of objects retrieved in limited time). While several methods are possible (from programming by hand to learning), the problem is made challenging as it is unlikely that one can predict the outcome of the interactions between robots without resorting to simulations, and because of the underlying complexity of addressing a decision problem with multiple agents with a sparse communication network.

During this internship, we will explore a class of distributed on-line learning algorithms for robotic swarms termed "embodied evolution" (Watson et al., 2002). In this context, each robotic agent follows a particular policy, which parameters can be sent and received to/from nearby agents. After a while, one of the received policy is selected depending on its measured performance and selection operator, then modified through mutation and/or cross-over, to be used as the new active policy for the agent. This results in an evolutionary-inspired algorithm that can be used to increase the group performance with respect to a particular objective, defined as an objective function to optimise (e.g. catch preys, retrieve objects, maximal area coverage, etc.).

One key problem is that, so far, the class of embodied evolutionary algorithms assumes that a local objective function is defined at the level of the individual. This a strong, and possibly wrong, assumption as the "true" contribution of one individual robot to the group may or may not depend on its particular performance. For example, let's imagine two scenarios: (1) if the goal is to retrieve different kind of objects, each of which requiring special behavioural skills, the global optimum is to converge towards two groups, each with particular functional specialisation ; (2) if the goal is to have the robots share resources in the environment, the global optimum is to converge towards individuals that retrieve as many resources as possible, as long as it does not steal resources from other robots (ie. the goal is not to maximise the number of resources retrieved per individual).

In order to address this problem, we will extend tools from cooperative game theory or recent approaches building synergy graphs, which makes it possible to estimate synergies and marginal contributions of individuals within a group (ie. how worth they are from the perspective of the whole) in order to drive adaptation towards a globally optimal collective behaviour. Due to the nature of the experimental setup, we will apply these tools under the constraint of distributed systems where the exact modality of communication from peer to peer may change over time (robots may get separated, or be densely packed together, depending on the task and the environment).

Practically, the internship will be composed of three main parts:

1. implementation of a basic embodied evolutionary algorithm on a group of 10 Thymio-2 robots (extended with a micro-computer and camera). In this scope, the local objective function will be a human-defined heuristic function, which provide a crude approximation of each robot performance wrt the group.
2. design and implementation of a method for estimating the marginal contribution of each robots (Shapley value, Banzhaf index) with global communication (simplification assumption)
3. extension of (2) to address setups where communication is restricted by the (spatial) distance between robots (ie. limited range).

## **Administrative information**

The internship will be done at UPMC, between the ISIR and LIP6 labs, starting Feb. 1st for 6 months (internship gratification is approx. 500 euros per month).

Co-advisors are: Nicolas Bredeche (ISIR/UPMC), Nicolas Maudet (LIP6/UPMC)

## **Short bibliography**

Embodied evolution: Distributing an evolutionary algorithm in a population of robots

RA Watson, SG Ficici, JB Pollack

Robotics and Autonomous Systems 39 (1), 1-18

Continuous Foraging and Information Gathering in a Multi-Agent Team

Somchaya Liemhetcharat, Rui Yan and Keng Peng Tee

Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), pages 1325-1333, May 2015

Weighted Synergy Graphs for Effective Team Formation with Heterogeneous Ad Hoc Agents

Somchaya Liemhetcharat and Manuela Veloso. Artificial Intelligence. Volume 208 (2014), pages 41-65, March 2014. DOI: 10.1016/j.artint.2013.12.002

A Review of Studies in Swarm Robotics

Bayindir, L., & Sahin, E. (2007) Turk J Elec Engin, 15(2), 115–147. <http://doi.org/10.1.1.98.7821>

Programmable self-assembly in a thousand-robot swarm

Rubenstein, M., Cornejo, A., & Nagpal, R. (2014) Science, 345(6198), 795–799. <http://doi.org/10.1126/science.1254295>

Designing collective behavior in a termite-inspired robot construction team

Werfel, J., Petersen, K., & Nagpal, R. (2014) Science (New York, N.Y.), 343(6172), 754–8. <http://doi.org/10.1126/science.1245842>