HAL
open science

# Learning by Imitation for the Improvement of the Individual and the Social Behaviors of Self-Organized Autonomous Agents

Abdelhak Chatty, Philippe Gaussier, Ilhem Kallel, Philippe Laroque, Adel M.Alimi

## To cite this version:

Abdelhak Chatty, Philippe Gaussier, Ilhem Kallel, Philippe Laroque, Adel M.Alimi. Learning by Imitation for the Improvement of the Individual and the Social Behaviors of Self-Organized Autonomous Agents. in IEEE International Conference on Swarm Intelligence (IC-SI)„ Jun 2013, China. pp.44-52. hal-00955890

**HAL Id: hal-00955890**

**https://hal.science/hal-00955890**

Submitted on 5 Mar 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Learning by Imitation for the Improvement of the Individual and the Social Behaviors of Self-Organized Autonomous Agents

Abdelhak Chatty[1,2], Philippe Gaussier[2], Ilhem Kallel[1], Philippe Laroque[2] and Adel M. Alimi[1]

[1] REGIM: REsearch Groups on Intelligent Machine, National School of Engineers (ENIS), Sfax University, Sfax, Tunisia
[2] ETIS: Neuro-cybernetic team, Image and signal Processing, National School of Electronics and its Applications (ENSEA), Cergy-Pontoise University, Paris, France

**Abstract.** This paper shows that learning by imitation leads to a positive effect not only in human behavior but also in the behavior of the autonomous agents (AA) in the field of self-organized creation deposits. Indeed, for each agent, the individual discoveries (i.e. goals) have an effect on the performance of the population level and therefore they induce a new learning capability at the individual level. Particularly, we show through a set of experiments that adding a simple imitation capability to our bio-inspired architecture allows increasing the ability of agents to share more information and improving the overall performance of the whole system. We will conclude with robotics' experiments which will feature how our approach applies accurately to real life environments.

**Key words:** Learning by Imitation; Cognitive Map; Emergent Structures.

## 1 Introduction

Swarm-based systems [1, 2] are now a classical approach to dealing with collective intelligence problems. In such approaches, to produce global emergent behaviors for Autonomous Agents (AA) the interaction between agents needs not to be complex [3, 4]. Based on this idea, researchers have been able to design a number of successful algorithms in the field of self-organized creation deposits. [5] has proposed a model relying on biologically plausible assumptions to account for the phenomenon of the clustering of dead bodies performed by ants. [6] showed that acting on objects simplifies the reasoning needed by a multi-agent system and allows the deposit of scattered objects. In the same field of creation deposits we assume an environment composed of several animats and three plants (A, B and C). The AA are motivated by the simulation of three types of needs related to the three plants and each need can be satisfied by a corresponding plant. The level of each type of need is internally represented by an essential variable, $e_i(t)$ whose value is in [0; 1] and varies with time $de_i/dt = -\alpha_n e_i(t)$, where $\alpha_n$ represents the decreasing rate of the essential variable. If a plant from the corresponding type has not been found, the agent dies. Thus, in order to maintain the

satisfaction level of our AA, keep them alive and optimize their planning, instead of only navigating between the three plants, it would be interesting if the agents were able to create relevant warehouses. So as to improve the performance of the system, several benefits can be expected from the imitation capability [7–9]. It can be considered as a powerful skill that would enable the AA to learn and discover new tasks and places. Learning by imitation is an intuitive and natural method, it is not only a skill useful for learning but also a way to speed up the learning process. Therefore, the researchers consider imitation as a powerful behavior which enables learning by observation even if the imitation was not intentional (i.e. imitation emerging from the ambiguity of the perception in a simple sensori-motor system) [7]. The aim of this work is to show the positive effect of the learning by imitation on the improvement of the performance of the deposits' system. To validate our system we performed a series of experiments with simulated agents and with swarm robots. The remainder of this paper is organized as follow: in section 2 the bio-inspired architecture is presented. Section 3 describes the behaviors of the AA. Sections 4 and 5 are devoted respectively to the description of the relevance of the emerging warehouses and to the explanation of the imitation process. Before concluding, section 6 and section 7 are concerned respectively with the analysis of the positive feedback induced by the imitation strategy in AA and in swarm robots.

## 2 The Bio-Inspired N.N Architecture of the Agent

Starting from neurobiological hypotheses on the role of the hippocampus in the spatial navigation, [10] produced a model of the cognitive map in the hippocampus representing the entire environment and not only the shortest paths to a given goal. The work of [11] revealed special cells in the rodents' hippocampus that strike off when the animal is at a precise location. These neurons have been called place cells (PC). In our model, we do not directly use PC we rather use neurons called transition cells (TC) [12]. A transition cell encodes for a spatio-temporal transition between two PCs consecutively winning the competition, respectively at time t and $t + \delta$t. The set of the PCs and the TCs constitute a non-cartesian cognitive map. A schematic view of the architecture of our AA is shown in fig. 1.

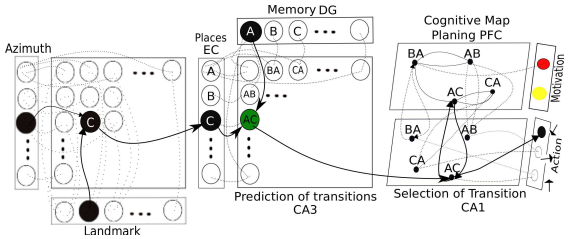


**Fig. 1.** Model of hippocampo-cortical for the building of an agent cognitive map

To create the PC, the agent takes, a visual panorama of the surrounding environment. The views are processed to extract visual landmarks. After learning these landmarks, a visual code is created by combining the landmarks of a panorama with their azimuth. This configuration serves as a code for PCs. The signals provided by the EC (the entorhinal cortex) are solely spatial and consistent with spatial cells activities. Spatial cells activities are submitted to a Winner-Take-All competition in order to only select the cell with the strongest response at a specific location. We will subsequently speak about the current location by indicating the spatial cell which has the highest activity at a given location. The temporal function at the level of the DG (dentate gyrus) is reduced to the memorization of a previous location. The acquired association at the level of CA3/CA1 (the pyramidal cells) is then the transition from a location to another in addition o the information concerning the time spent on carrying out this transition. Once the association from the previous location and the new one is learned, every new entry will reactivate the corresponding memory in the DG. During exploration of the environment, the cognitive map is gradually created as the agent moves (see fig. 2).

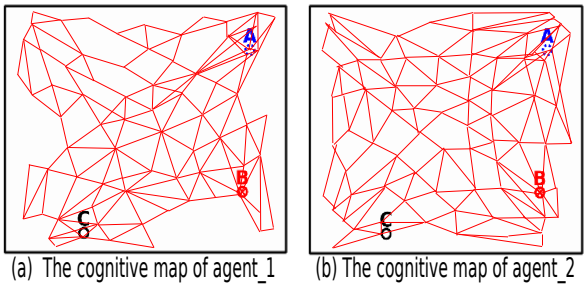


**Fig. 2.** The cognitive map of two agents at 5104 time steps. The shape of the cognitive map in (a) and (b) proves that its construction is related to the agent's own perception.

## 3 The Behaviors of the Autonomous Agents

In the context of situated cognition, local rules can lead to create emergent structures allowing the creation of warehouses relevant to the sorting strategy used by [5]. We propose local generic rules depending on the number of withdrawals and deposits of warehouses according to the number of agents perceived. The agent is indeed ought to favor the creation of warehouses in locations which contain other agents rather than empty regions. Thus, the perception of local agents is responsible for controlling the rules of withdrawals and deposits. The condition for withdrawing is computed by $Pr_{(Taken)} = \exp^{-\lambda N_A}$ where $N_A$ is the number of agents in the neighborhood, $\lambda$ is a positive constant: The probability that an agent needs to take some plant goods increases when it perceives that the

plant is less used by other agents. So, the more agents are near a plant, the less withdrawal there is and vice versa. The condition for deposits is computed by $Pr_{(Deposits)} = (1 - \exp^{-\alpha N_A}) * (1 - \exp^{-\beta t})$ where $\alpha, \beta$ are environmental factors, $N_A$ is the number of agents in the neighborhood and t is the time since the taking: the probability of deposit increases with time and distance from the origin plant (when the agent is far from the origin plant where it took the last goods) and it depends on the number of the agents in the neighborhood (when the current place of the agent is frequented by other agents). For this reason, we tried to restrict the ability of agents to perceive the environment. Moreover, the deposit operation is also built on the concept of refueling : the agent puts goods in the warehouses that already exist.

## 4 The Emergence of Warehouses

We use an environment with 60 agents (under 20 agents, the system will not be able to create a fixed number of warehouses in fixed places) and three original plants. This environment is continuous and in order to cross it diagonally, an agent needs 200 time steps. This experiment is just a general scenario to describe the behavior of the agents and the creation of relevant warehouses.

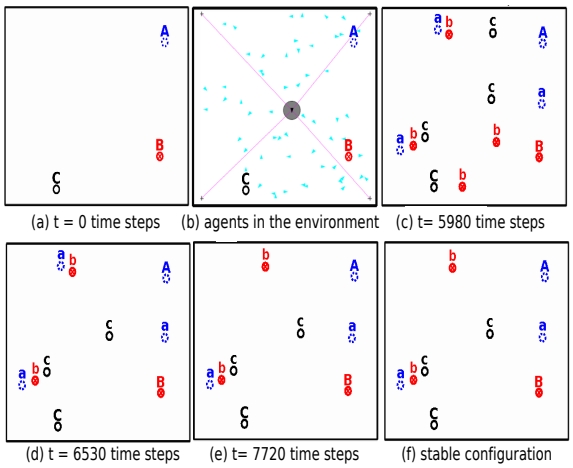


**Fig. 3.** The figure shows an example with 3 plants A, B and C whose positions are fixed and can deliver an unlimited amount of products. They provide warehouses of 3 different kinds ("a" is a warehouse-type A, "b" is a warehouse-type B and "c" is a warehouse-type C). Through the individual deposit process of agents, warehouses emerge and become stable after a while.

The agents start to move randomly in the environment, with a limited range of visual perception (see Fig. 3b). While passing through a plant, an agent increases

its level of satisfaction and applies the local rules of taking and transporting a quantity from the associated product. If the decision of transporting some products is taken, the agent continues its journey. If deposited, goods represent a new warehouse (which allows 10 visits for agents) permitting others to increase their satisfaction level. Agents also have the possibility of refueling warehouses by adding products to them (the available products in the warehouses will increase). This provides stability for warehouses in relevant locations which are close to several agents in order to prevent loss. However, warehouses which are abandoned or poorly visited will eventually disappear since the amount of goods available will decrease rapidly. Fig. 3c and d show the disappearance of isolated warehouses. When a planning agent tries to reach a warehouse and realizes that it has disappeared, the agent dissociates the current PC from the formerly-corresponding warehouse and resets the motivation to 0. Hence the PC does not fire any more when the agent feels the need for this warehouse. Similarly, when a new matching warehouse is discovered, the paths leading to the warehouse are immediately reinforced, modifying the cognitive map synchronously with the environment. In fig. 3e the AA converges to a stable solution with a fixed number of warehouses in fixed places at 7720 time steps and remains the same for more than 20000 time steps (see fig. 3f).

## 5 The Imitation Process

Collective learning experience gathered by individual agents cannot be of use to others, if there is no means to communicate information between agents. One way to add knowledge transmission among agents is to make them able to imitate one another. Thus, at the individual level, agents can rely on an on-line, continuous building of a cognitive map whose structure depends on their own experience and discovery of the environment; and at the social level, they can take advantage of the ability to imitate one another using simple agent-following strategies to transmit parts of one agent's cognitive map to another's, leading to some kind of naturally distributed knowledge, similar to what can be achieved in swarm-intelligence systems, except that shared knowledge does not use the physical space as repository but uses individual cognitive maps instead. We implemented and studied a simple imitation strategy based on the azimuth: if a single agent is visible, it becomes the chosen imitation target; if several agents are visible, then the chosen target is the closest one to the direction of the agent which tries to imitate. The decision of imitation is controlled by a probabilistic function (see eq. 1): the probability of imitation decreases when the number of discovering of different kind of resources(plants/warehouses) $N_D$ increases.

$$Pr_{(Imitation)} = \eta \exp^{-\rho N_D} \tag{1}$$

$\eta$ and $\rho$ are positive values. To better understand the imitation behavior of agents, fig. 4a shows that the agents represented by triangles, started to move randomly in the environment. Due to the imitation behavior fig. 4b shows that subgroups of agents appeared in the environment. According to the curve (d),

this phase coincides with an increase in the number of agents in imitation. Indeed, the agents try to follow each other so as to localize and learn the resources' positions. We can also see that some subgroups, have succeeded to discover the plant A and B and to learn their positions. After discovering all kinds of resources (where the average time is 505 time steps. However without imitation, the average time to reach the resources is 1683 time steps) the probability of the imitation will decrease (see fig. 4c). Thus, the number of agents in imitation shown by the curve (d) also decreases, since agents are now able to return to the resources on their own.

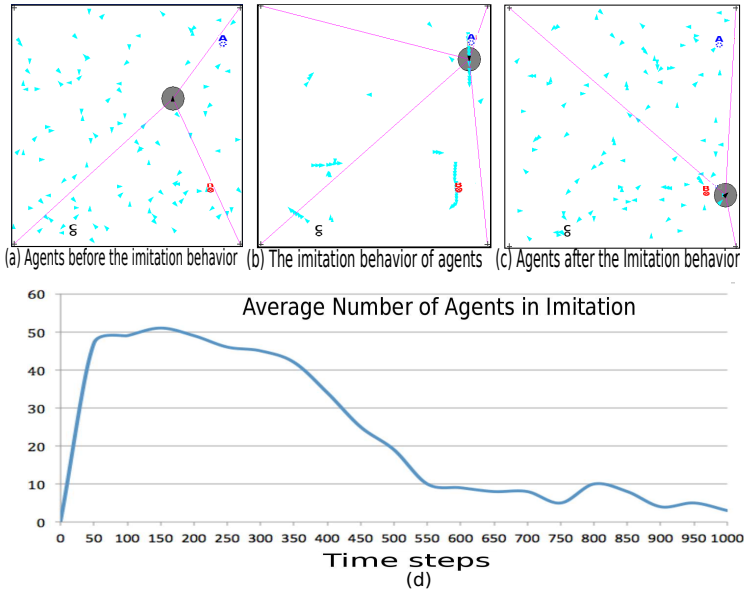


**Fig. 4.** The evolution of the agents' number in imitation through 1000 time steps.

## 6 The Positive Feedback of the Imitation Behavior

The behaviors of the AA are not deterministic because they do not always provide the same results on several tests with the same parameters. In order to obtain complete results, we kept the same previous parameters ($\alpha = \beta = 0.3, \lambda = 0.7, \eta = 0.1, \rho = 3, ST = 50$ and the number of agents is 60) and we conducted ten tests of AA for each experiment until 20000 time steps (the number of tests is determined by the statistical Fisher test). Through the imitation behavior we show that the agents can more rapidly (i) learn the position of the resources (ii) create a relevant warehouse and (iii) satisfy their needs. Thus, the agents would be able to improve the performance of the whole system.

### 6.1 Optimization of Warehouses' Numbers and Convergence Time

We started to experimentally keep track of the number of warehouses' visits when the agents imitate as compared to the number of warehouses' visits when the agents do not have imitation capabilities (for 20000 time steps). The average number of visits to warehouses in the first case (246) is more important than in the second one (115). This shows that with the imitation behavior, the warehouses are more frequented by the agents. We defined the performance of the deposit system through the convergence time and the number of warehouses in the environment. The analysis of the agents' behavior in Table 1 enabled us to show that due to the imitation behavior the AA were able to deposit a fixed number of warehouses which can emerge in fixed positions while minimizing the warehouses numbers and the convergence time without having to use thresholds in order to limit the number of warehouses nor to specify their locations.

**Table 1.** Average of warehouses number and convergence time.

| | Warehouses Number | Convergence Time |
|---|---|---|
| Without Imitation Behavior | 7 | 9000 time steps |
| With Imitation Behavior | 6 | 8123 time steps |

### 6.2 Optimization of Agents' Planning and their Survival Rate.

The AA were also able to create emergent and stable warehouses allowing an optimization of the planning. Table 2 shows that due to the imitation, agents can optimize their planning time with the help of relevant warehouses. This leads to a higher average of satisfaction level. Indeed, based on the imitation, the AA is able to better optimize the planning time of agents thanks to the faster learning of the new places of warehouses and to improve their level of satisfaction. To study the influence of the imitation strategy on the survival rate of the populations, we used the same environment and launched 20, 30, 40 then 60 agents, and counted the number of agents that survived, or died for not having found all of the three types of resource. The results show that adding an imitation capability can dramatically enhance the survival rate of the population from 45,66% (without imitation behavior) to 60,64% (with imitation). Indeed, coupling imitation behavior with the cognitive map allows agents to discover and to learn the position of the resources in the environment more rapidly

**Table 2.** Optimization of planning time.

| | Without Imitation Behavior | With Imitation Behavior |
|---|---|---|
| Planning Time | 450 time steps | 218 time steps |
| Satisfaction Level | 88.07 | 91.25 |

## 7 The Effect of Imitation Behavior on Swarm Robots

We have also validated the positive feedback of the imitation strategy on a minimal robotics setup [13] based on the same bio-inspired architecture which has been validated in real robots in [14] . We conducted 20 experiments (this number is determined by the statistical Fisher test) composed of an imitator robot $IR$ (having an imitation behavior) and with a leader robot $LR$ (that has already learned the position of the goals G1 and G2 with a threshold of vigilance to learn new places equal to 0.65 and a duration of learning equal to 30 minutes). Fig. 5a, b, d and e show that the $IR$ tries to follow the $LR$ which is looking for the goals. Fig. 5c and f demonstrate that the $IR$ succeeded in discovering both goals during the following of the $LR$. As in the experiments with AA, this experience allows us to show through imitation that the $IR$ was able to find both goals and to optimize the time of their discovery. Indeed, the $IR$ takes 5 minutes to find both goals. However, a robot moving randomly (without an imitation behavior) takes 22 minutes (about 4 times more) to discover the two goals.

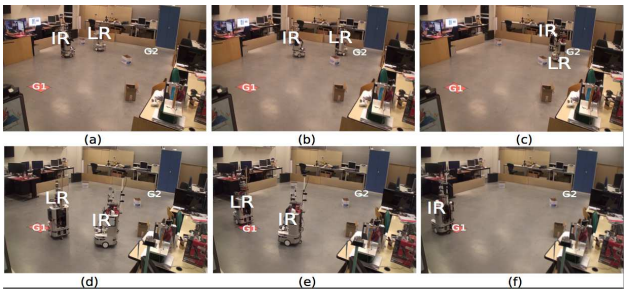


**Fig. 5.** The influence of the imitation behavior in swarm robots.

## 8 Conclusion

In this paper we have tried to show the importance of the learning by imitation which leads to the improvement of the performance of the whole system. As a possible application, we have added in [15] customers to the environment in order to resolve the classical warehouse location problem. We also highlighted the importance of the imitation strategy in real robots which allows to solve the navigation task and optimizes the time to explore the plants. As prospects, we try to study the limits of emergent structures in the real world.

## Acknowledgment

The authors would like to thank the financial support of the Tunisian General Direction of Scientific Research and Technological Renovation (DGRSRT), under the ARUB

program 01/UR/11 02, the Institut Français de Tunisie and the NEUROBOT french project referenced ANR-BLAN-SIMI2-L2-100617-13-01.

## References

1. Bonabeau, E., Theraulaz, G.: Intelligence Collective. Hermes (1994)
2. Brooks, R.A.: Coherent behavior from many adaptive processes. In: Proceedings of the third international conference on Simulation of adaptive behavior, Cambridge, MA, USA, MIT Press (1994) 22–29
3. Mataric, M.J.: Designing Emergent Behaviors: From Local Interactions to Collective Intelligence. In Meyer, J.A., R.H., Wilson, S., eds.: Proceedings of the Second Conference on Simulation of Adaptive Behavior, MIT Press (1992) 1–6
4. Chatty, A., Kallel, I., Gaussier, P., Alimi, A.: Emergent complex behaviors for swarm robotic systems by local rules. In: IEEE Symposium on Computational Intelligence on Robotic Intelligence In Informationally Structured Space (RiiSS). (April 2011) 69 –76
5. Deneubourg, J.L., Goss, S., Franks, N., Franks, A.S., Detrain, C., Chrétien, L.: The dynamics of collective sorting robot-like ants and ant-like robots. In: Proceedings of the first international conference on simulation of adaptive behavior on From animals to animats, Cambridge, MA, USA, MIT Press (1990) 356–363
6. Gaussier, P., Zrehen, S.: Avoiding the world model trap: An acting robot does not need to be so smart! Robotics and Computer-Integrated Manufacturing **11**(4) (1994) 279 – 286
7. Gaussier, P., Moga, S., Banquet, J.P., Quoy, M., Modelisations, N.E.: From perception-action loops to imitation processes: A bottom-up approach of learning by imitation (1997)
8. Schaal, S., Peters, J., Nakanishi, J., Ijspeert, A.: control, planning, learning, and imitation with dynamic movement primitives. In: workshop on bilateral paradigms on humans and humanoids, ieee international conference on intelligent robots and systems (iros 2003). (2003)
9. Chaminade, T., Oztop, E., Cheng, G., Kawato, M.: From self-observation to imitation: visuomotor association on a robotic hand. Brain research bulletin **75**(6) (April 2008) 775–784
10. Muller, R.U., Stead, M., Pach, J.: The hippocampus as a cognitive graph (1996)
11. O'Keefe, J., Nadel, L.: The hippocampus as a cognitive map / John O'Keefe and Lynn Nadel. Clarendon Press ; Oxford University Press, Oxford (1978)
12. Gaussier, P., Revel, A., Banquet, J.P., Babeau, V.: From view cells and place cells to cognitive map learning: processing stages of the hippocampal system. Biological Cybernetics **86**(1) (2002) 15–28
13. Chatty, A., Hasnain, S., Gaussier, P., Kallel, I., Laroque, P., Alimi, A.: Effect of low level imitation strategy on an autonomous multi-robot system using on-line learning for cognitive map building. In: IEEE International Conference on Robotics and Biomimetics (ROBIO). (December 2012) 1–6
14. Chatty, A., Gaussier, P., Kallel, I., Laroque, P., Alimi, A.: Adaptation capability of cognitive map improves behaviors of social robots. In: IEEE Conference on Development and Learning and the Epigenetic Robotics. (November 2012) 1–6
15. Chatty, A., Gaussier, P., Kallel, I., Laroque, Florance, P., P., Alimi, A.: The evaluation of emergent structures in a cognitive multi-agent system based on online building and learning of a cognitive map. In: 5th International Conference on Agents and Artificial Intelligence (ICAART). (February 2013) 269–275