

Impact of imitation on the dynamics of animat populations in a spatial cognition task

Ph. Laroque*, N. Cuperlier, Ph. Gaussier

Neurocybernetic team, ETIS, CNRS (UMR 8051) / ENSEA / UCP

Universit de Cergy-Pontoise, 2 rue Adolphe Chauvin, 95302 Cergy-Pontoise cedex

Abstract This paper focuses on the dynamics that can emerge from animat populations [7] when the individuals can only use single sensori-motor learning (reactive behaviour) or can learn a map from their environment (cognitive behaviour). Animats have to find their way to resources that can help them to survive. Adding an – even simple – mimicing mechanism allows for newly incoming animats to take benefit of the experience of their elders to find resources more quickly. Our experiments show that the level of knowledge of the animats has an important and different impact on the effect of the imitation ability itself; we show that the complexity of the environment also has an important impact on this effect. That leads us to think that some kind of meta-control can be necessary to let the animat choose a suboptimal strategy whenever it gives better performance, namely in some kinds of simple environments

Keywords: Imitation; Learning; Neural networks; Planning; Cognitive map; Social interaction; Collective intelligence.

1 Introduction

The capability to solve a problem using a population of agents has been intensively studied for more than one decade. Most of these works focus on insect-like agents using reactive strategies [11] and leaving information (such as pheromons) in their environment to allow a collective decision taking. Situated agents or animats [7] functioning on the basis of local and incomplete informations have been shown to be very efficient on a wide variety of tasks. In our laboratory, we have shown, for instance, that not so complex animats can learn to solve complex choice problems like going to drink to a farther source because they can also find food that will be important later when their internal motivation (or goal) will change. Our solution is based on the building of a cognitive map of the environment and the use of an on-line hebbian reinforcement of the most used paths on the map [8].

However, very few works have been done on the comparison between the global capabilities of a population of more or less complex animats in a situation with or without social interactions. The work presented here is part of a CNRS project “Geomatics, Space, Territories and Mobilities”. In collaboration with a team in spatial economics, our long term goal is

*Corresponding author.

E-mail address: Philippe.Laroque@dept-info.u-cergy.fr

to build a taxonomy of the different kinds of spatio temporal dynamics that can emerge from a population of agents according to their cognitive capabilities, their imitation or interaction capabilities and the complexity of their environment. It is clear that the criteria that can be used for such a taxonomy (for instance, based upon the complexity of the environment, or upon the agents' imitation or adaptation capability) need tests to be discovered and trusted. In this paper, we will try to compare a simple reactive control architecture (using only simple associative sensori-motor learning) with a more complex architecture allowing an animat to learn and use a cognitive map of its environment. We will confirm some well-known intuitions on the interest of imitation [3, 4, 6] and present and discuss some surprising results showing a complex dependency between the learning capabilities of the animats and the complexity of the environment when a simple imitation mechanism is introduced. We stress here on trying to quantify the impact of imitation as those parameters evolve. That is the reason why the results presented here are given for each possible "intelligence level / imitation ability" pair.

Imitation processes are usually divided in two levels: the *action level* of imitation [1] is related to the mechanisms involved when reproducing a simple action, often an elementary movement. The *program-level* refers to imitation of complex actions while preserving their organisational level. The work presented in this paper mainly focuses on the first of those two levels. Our goal here is to demonstrate that a smarter global behaviour does not necessarily depend on the individual planning capacity (as is the case for cognitive animats [8]), since even purely reactive autonomous animats can take benefits of imitation mechanisms. The fact that imitation can help to bring social interactions among animats, leading to peculiar local dynamics that can be observed, is also of importance: for instance, the ability for a given generation of animats to find stable states more easily, as in [10], is indeed the condition for such a generation to survive.

2 Material, model and method

In our experiments, the animats are living in an unknown environment made of resources and obstacles. They only receive two types of information, the points of interest (landmarks, obstacles, sources, other animats) they can see from their current position, and the azimuth under which they perceive every remarkable point.

The visual aspect of the environment is shown in Figure 1. The animats need to visit three kinds of resources (food, water and nest) to survive: these can make them periodically bring their vital satisfaction levels (hunger, thirst and stress) back to a correct, sustainable value, which decreases over time. To reach the different sources, the animats have to avoid obstacles through which they can't go. They can only see things that are within a given perimeter, except for landmarks which are always visible, unless occluded by some obstacles. The animats have to switch periodically between several strategies, according to their internal state and level of cognition.

We decided to study reactive versus cognitive animats. Both species have in common the use of *place cells* [8, 9] to record newly discovered positions in the environment and be able to come back to interesting locations when needed. Place cells are only created from visual information (*landmark – azimuth* couples), so there is no need for the animats to have access to any kind of coordinates. The differential equations that rule the creation of place cells, and those managing the satisfaction level of the animat over time can be found in [8] (see also[5]) – both papers are available at <http://www.etis.ensea.fr>

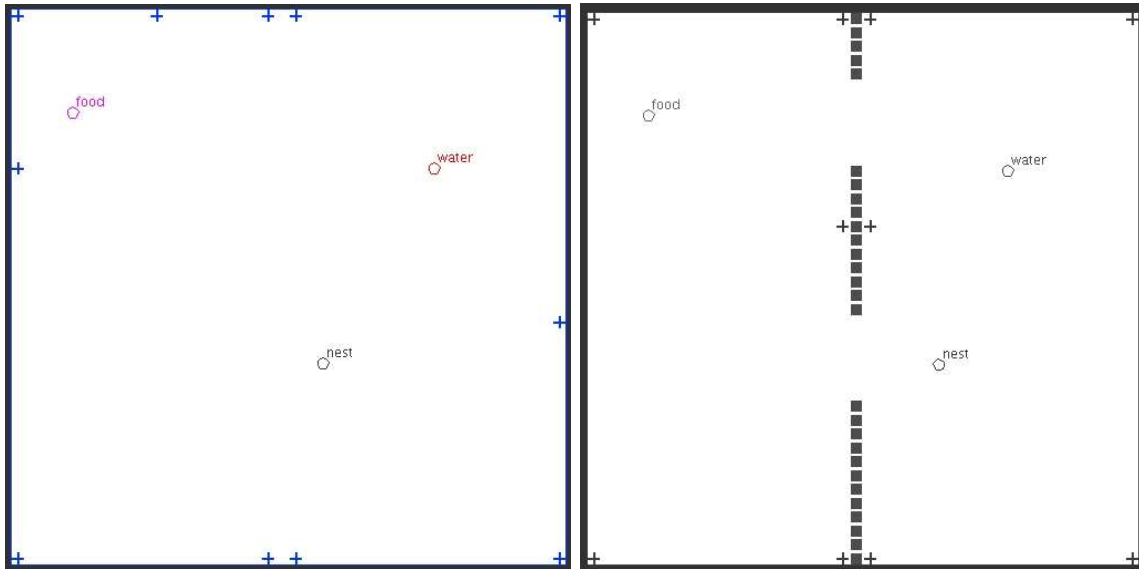


Figure 1: The animats in their environment, without obstacles (*left*) and with several walls (*right*).

At start, the animat does not know anything about its environment, except that there are some landmarks. During its moves, it receives the perception (a list of *landmark – azimuth* couples) of its current position on the map. A distance with previously learned positions is then computed, and if below a certain threshold, the new position is recorded on a new neuron. When a resource is discovered, the associated place cell records the presence of the resource; that allows for activating this place cell when the need for the corresponding resource arises. The animat must discover one source of each type (food, water and nest) to survive. Once the three kinds of sources have been discovered, a stable state can be reached in simple cases: the animat can now live forever, since resources are not limited for the time being. We will show that, when the complexity of the environment increases, simple (reactive) agents can not reach such a stable state – even if they discover the three sources – whereas the performance of cognitive agents increases dramatically.

Possible strategies for purely reactive animats

A reactive animat can be in one of the following states, in decreasing precedence order: (i) obstacle avoidance; the presence of an obstacle makes the animat deviate from the direction it intended to follow, (ii) attraction by one of the (previously discovered) sources; this occurs when the satisfaction level of one of the animat's vital variables goes below a given threshold, when the source is easily reachable (*e.g.*, there is no obstacle between the animat and the source and the environment is static) and the animat knows where it can find it, (iii) imitation; this occurs when an animat detects another one and decides to follow it; it is one way to quickly discover new resources (if the followed animat already knows some of the sources, it can bring it to them), and (iv) random exploration; this occurs when all vital motivations are satisfied and either the animat did not see any other animat, or it decided not to imitate the ones it saw, or it does not know where to go. It is the second way the animat can discover the existing resources. This set of abilities and constraints (visual navigation, simple imitation mechanism) has been successfully tested for years on actual robots in our laboratory [5].

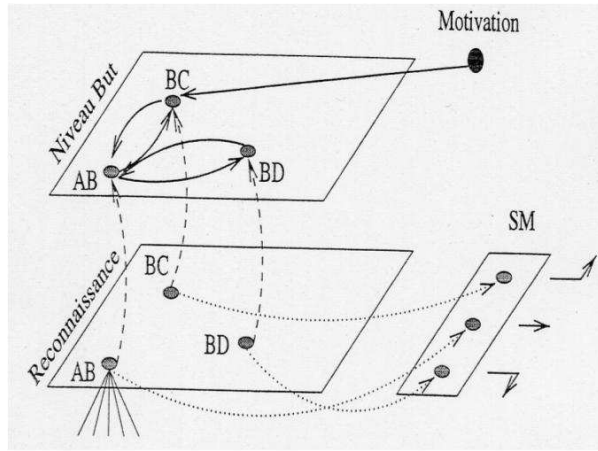


Figure 2: Basic principle of a cognitive map to control animat navigation from the recognition of particular transitions[5].

Possible strategies for cognitive animats

The material briefly described above is sufficient for reaching a learned goal, if each place cell is associated with a movement to perform. Nonetheless, some important problems can't be tackled by such a simple architecture, notably when the goals do not all belong to the same visual environment (no visual connectivity). To address these problems, cognitive animats can rely on a cognitive map [8, 9] to come back to a previously discovered interesting point, even in an evolving environment. Such a map is built by linking together two place cells reached successively (see figure 2). That way, exploring the environment leads from an initially random map to accurate connections between place cells. The priority of behaviours is the same as for reactive animats, except that planning strategy replaces a simple gradient-following strategy used to reach the desired sources. This technique is distinct from other strategies such as Q-learning [13] since they are generally unable to reuse what has been learned from one motivation to another. Moreover, the state space in which animats live is a continuum, which would lead to a huge number of states.

Rules for starting / stopping imitation

Each animat can choose to follow another animat as soon as it becomes visible. Animats are undistinguishable from one another: when a "teacher" (imitated) animat gets out of sight, the "student" (imitating) animat has no way to remember who it was and find it again. The decision to start imitating is taken according to a random value being over or under a given, constant threshold. The choice of who to follow is ruled as follows: if a single animat is visible, it becomes the chosen imitation target; if several animats are visible, then the chosen target is the one that is closest to the direction in which the animat is currently running (that is, the one whose perceived azimuth is closest to $0 \text{ mod } 2\pi$).

The main problem we encountered with this simple imitating scheme is deadlocks due to reciprocal imitations. When two animats came in front of one another, they sometimes both decided to imitate each other, leading to a blocking situation: each of them decided to follow the other, so they changed their direction at the same time and in the opposite direction!

In order to avoid / lessen reciprocal imitation decisions that lead to such deadlocks, only animats whose azimuth is less than or equal to $\pi/2$ (in absolute value) are competing. Such problems are an illustration of the kind of local dynamics that can emerge from even simple social mechanisms (here, the creation of a moving loop of animats following each other until death). When the target teacher gets out of sight, the student stops imitating but the process can start again later. For the time being, no feedback on the benefits and/or drawbacks of imitating is used to adapt the imitation decision condition.

It would have been possible to completely solve this problem differently, for instance by differentiating the animats so that one animat can know who it is imitating (namely, an elder); however our goal in this first step is to keep the learning mechanism as simple as possible, in order to be able to interpret results without any bias. Yet another technique is to allow animats to “decide” to stop imitating when they receive no reward after a certain amount of time.

3 Experiments and results

This first series of experiments aims at studying three aspects of the influence of imitation: (i) the survival rate of the animats, (ii) the need for a more or less complete exploration of the map to find all resources, and (iii) the influence of *detour* problems. Each of those aspects is illustrated with an experiment involving reactive and cognitive animats, with and without imitation: every possible combination is tested.

3.1 Survival rate of a generation of animats

We sequentially launch five generations of ten animats, waiting for stabilization of generation i to launch generation $i + 1$. The map is paved with 40x40 square regions (sources and landmarks occupy one single region). After a certain amount of time (directly depending on essential variables thresholds), only those of the animats that found the three sources survive. Then a new generation is launched and join the survivors of the previous ones; its members can, if imitation is turned on, take benefit of the fact that the elders (who can be viewed as “teachers” to a certain extent) know where the sources are to find them more quickly, namely before their essential variables fall to zero. Each experiment has been tested 20 times, each time with the same initial conditions, but different series of random numbers to drive the animats exploration phases. The average results are presented in figure 3. We can see that all curves are roughly linear, and that the results concerning animats capable of imitating (upper curve in both cases) are better than those for non-gregarious animats. However, what seems more surprising is the fact that, while reactive animats take high benefits of the growing number of “teachers” (the upper curve of the left diagram is ascendent), it is not the case for cognitive animats, for whom the proportion of potential teachers does not seem to have a significant impact. The return on investment here is particularly low, since the used algorithm is much more complex than in the reactive case. This tends to prove that learning and using a complex coding of the environment may not be efficient in a very simple environment. Indeed, cognitive animats learn and use an incomplete map of their environment, so their behaviour can be less optimal than that of reactive animats. Even more when imitation ability is turned on: reactive animats end at learning an optimal behaviour in an open environment, while cognitive animats go on using complex, suboptimal solutions discovered during their first free exploration. The suboptimality will be more important if the motivations create a

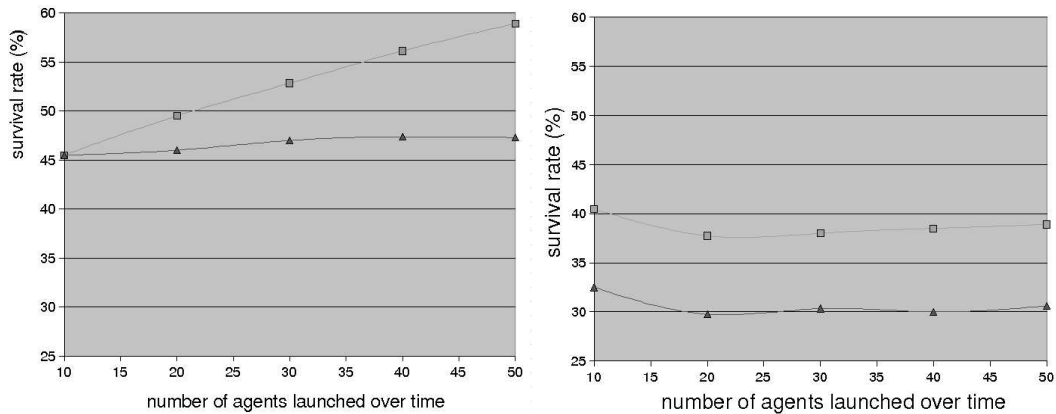


Figure 3: Experiment 1: survival rate of a generation of reactive (*left*) and cognitive (*right*) animats. Squared curves summarize results with imitation turned on, triangles are used when imitation is turned off.

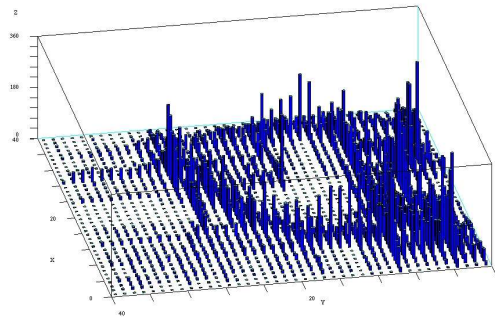


Figure 4: example of map exploration rate

high level of pressure on the action selection mechanism of the animat: it can lack the time needed for further exploration, needing to always go from one source to another one.

These results are encouraging and seem to confirm our hypothesis. However, we must bear in mind that the number of surviving animats varies and that the results only reflect an average behaviour. Moreover, they are highly dependent on the value of the essential variables.

3.2 Exploration rate

Here, we are interested in studying the average portion of the map explored by animats, and the average time it takes them to reach a stable state (*i.e.* to learn where the three sources are located). The map is now paved with 60x60 square regions, so it is 2.25 times bigger than for the first experiment. Figure 4 shows a typical repartition of the explored regions.

We first launch as many animats as needed to have ten of them knowing all sources, thus acting as potential “teachers”. We inhibited imitation for those teachers. Then we launch new animats (*students*), one at a time, to make sure that an animat always imitates a teacher (and not another student), until ten of them survive. In this test, students have a 360-wide vision, since we don’t have to deal with problems such as reciprocal imitation. We also increased

	Stabilization time		portion of map explored	
	\bar{t}	$\sqrt{\sigma}$	% explored	$\sqrt{\sigma}$
reactive, non-imitating	417.3	56.56	6.46%	1.01
cognitive, non-imitating	441.43	42.14	6.74%	0.64
reactive, imitating	189.46	37.05	3.41%	0.6
cognitive, imitating	240.22	34.96	3.67%	0.49

Figure 5: Experiment 2: stabilization time and exploration rate for cognitive and reactive animats.

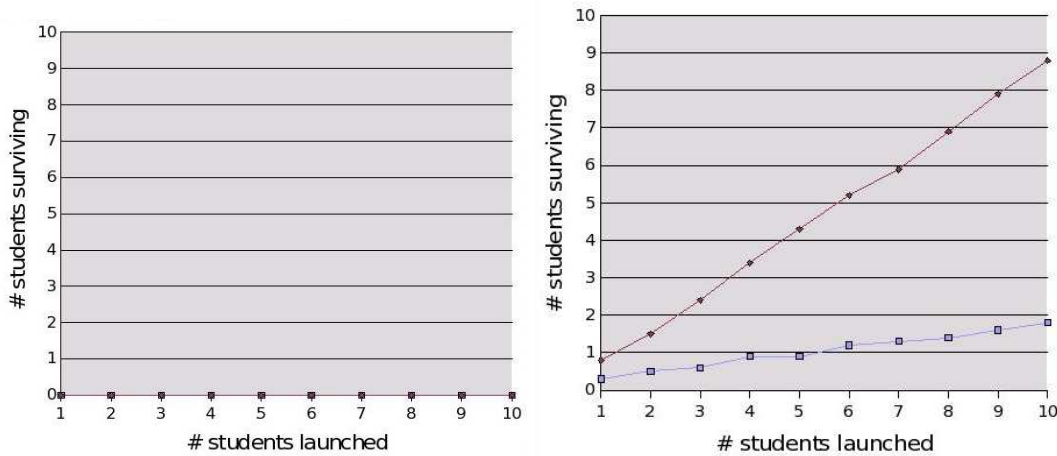


Figure 6: Experiment # 3: influence of obstacles on reactive (left) and cognitive (right) animats. Gradient descent does not allow reactive animats to correctly navigate between sources.

the lower level of the comfort area associated to each essential variable from 40% to 70%, in order to let the animats have to come back more frequently to the different sources. The test has been iterated 15 times and the average results are presented in figure 5.

These results show that students explore a much smaller part of the environment and – which is closely related – that they need less time than teachers to learn the location of the different sources. Reactive and cognitive animats reveal roughly similar performances for this test, which is not that surprising: the effect of imitation is maximum on this experiment and flattens the difference between both types of agents.

3.3 Influence of environment complexity

In order to highlight the differences between reactive and cognitive animats, the third tests run in an environment with obstacles (see figure 1, right part). First, ten “teachers” are launched, one after another. Only surviving teachers are taken into account. Then students are launched, under the same conditions as for experiment #2. The results are presented in figure 6. The presence of obstacles prevents reactive animats to rely on the same set of landmarks on the whole map, so – without any possibility to remember the path that eventually led them once to a source – their decisions about the direction to follow to reach back the sources are almost always wrong (no global attraction basin). As seen on the left diagram, reactive animats never survive. As far as cognitive animats are concerned, however, we can

see that releasing the limitation on the angle of vision of the animats greatly improves the performance of imitation (upper curve), compared to the results of experiment #1.

4 Conclusion and perspectives

This paper is a first step of a multi-disciplinary approach about studying and comparing behaviours of populations of more or less complex architecture to control inline learning on software agents or robots. Our long-time goal is (i) to better understand the underlying control strategies, and (ii) to highlight the emerging links between problem specialization and cognitive ability about the animats. For instance, it will be interesting to duplicate sources of the same nature and observe the potential formation and of subgroups, particularly as sources become unavailable after a certain time.

We believe – and the first results reported here seem to legitimate our view – that new incoming animats will naturally tend to follow one of the optimal solutions found by their elders, implementing a simple kind of global memory. In that sense, such animat populations form some kind of auto-poietic system [12]. However, the results shown here plead for the need to merge information coming from purely sensori-motor sources and from “intelligent” action selection mechanisms at a higher level: to handle this problem, we probably need to implement some kind of emotional meta-controller [2].

Acknowledgments

The *Geomatics* project is part of the *Information Society* CNRS Program. The project members are IERSO (Institut d’Economie Rgionale du Sud-Ouest, IFRReDE EA2956), ETIS (Equipe Traitement de l’Image et du Signal, UPRESA 8051) and ENST-Bretagne.

References

- [1] P. Andry, P. Gaussier, J. Nadel, From visuo-motor Development to low-level imitation, 2nd International Workshop on Learning Robots-EWLR’98, Edinburgh, UK, 1998.
- [2] Avila-Garca, O., Caamero, L. 2002. A Comparison of Behavior Selection Architectures Using Viability Indicators. In Proc. of the EPSRC/BBSRC International Workshop Biologically-Inspired Robotics: The Legacy of W. Grey Walter, pp. 86-93. August 14-16, 2002, HP Labs Bristol, UK.
- [3] Aude Billard and Maja J Mataric, Learning human arm movements by imitation: Evaluation of a biologically inspired connectionist architecture, *Robotics and Autonomous Systems*, 37:2-3, Nov 30, 2001, 145-160.
- [4] K. Dautenham, C.L. Nehaniv, *Imitation on Animals and Artifacts*, Cambridge, Mass., USA, MIT Press, 2002.
- [5] P. Gaussier, C. Joulain, S. Zrehen, J.P. Blanquet, A. Revel, Visual navigation in an open environment without map, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS’97)*, Grenoble, 1997, pp 545-550.
- [6] Y. Kuniyoshi, The science of imitation - towards physically and socially grounded intelligence. Special issue TR-94001, Real World Computing Project Joint Symposium, Tsukuba-shi, Ibaraki-ken, 1994.
- [7] Jean-Arcady Meyer, *The Animat Approach: Simulation of Adaptive Behavior in Animals and Robots*”, NSI98

- [8] M. Quoy, P. Laroque, P. Gaussier, Learning and motivational couplings promote smarter behaviors of an animat in an unknown world, in: *Robotics and Autonomous Systems*, vol. 38, Numbers 3-4, march 2002, pp. 149-156.
- [9] A. Revel, P. Gaussier, J.P. Blanquet, Taking inspiration from the hippocampus can help solving robotics problems, in: *European Symposium on Artificial Neural Networks*, Bruges, Belgium, IEEE Press, New-York, 1999.
- [10] L. Steels: *The Talking Heads Experiment. Volume 1. Words and Meanings*. Antwerpen, 1999.
- [11] Theraulaz G., Goss S., Gervet J. & Deneubourg J.-L. (1991). Task differentiation in *Polistes* wasp colonies: a model for self-organizing groups of robots. In: *From Animals to Animats, Proc. of the 1st Intl. Conf. on Simulation of Adaptive Behavior* (Meyer, J.A. & Wilson, S.W., eds), 346-355, MIT Press.
- [12] F. Varela, *Principles of Biological Autonomy*, New York: Elsevier (North Holland), 1979.
- [13] C.J.C.H. Watkins, *Learning from delayed rewards*, PhD Thesis, Psychology Dept, Cambridge University, England, 1989.