Guiding for Associative Learning : How to Shape Artificial Dynamic Cognition ?

Kristen Manac'H, Pierre de Loor

Guiding for Associative Learning : How to Shape Artificial Dynamic Cognition

Kristen Manac'h and Pierre De Loor

UEB – ENIB-CERV – LISyC
Brest – France
Email: manach,deloor@enib.fr

Abstract. This paper describes an evolutionary robotics experiment, which aims at showing the possibility of learning by guidance in a dynamic cognition perspective. Our model relies on Continuous Time Recurrent Neural Networks and Hebbian plasticity. The agents have the ability to be guided by stimuli and we study the influence of a guidance on their external behavior and internal dynamic when faced with other stimuli. The article develops the experiment and presents some results on the dynamic of the systems.

1 Positioning

Works in cognitive science show that cognition comes from the interaction between brain, body and environment [14, 12, 1]. Then, a cognitive system can be considered as an autonomous complex system disturbed by the environment, whose representation lies in sensory-motor invariants. This leads to different theoretical proposals in robotics and in artificial life [8, 2, 4]. Following these perspectives, some applications propose the use of dynamic systems [10, 9, 7]. Through interactions, these systems evolve between attractors and present plenty of different possible evolutions. These dynamic properties are the basis for adaptation, decision, memorization and also creation processes [13]. Different works address the sensory-motor invariants acquisition in such systems by pregiven adaptive behavior using Continuous Time Recurrent Neural Networks (CTRNN) which can approximate any dynamic system [5]. For example, [3] build photo-taxic robots which can adapt to sensors inversion at an ontogenetic scale.

We address the problem of learning behavior for such a system thanks to a specific interactive loop : a guidance signal. Indeed, learning by guidance is a more complex kind of adaptation because it can lead to different final behaviors. It is characterized by irreversibility which is a crucial property for ontogeny. Consequently, our concern is to establish experiments which will enable us to move from the status of self-adaptation to that of evolution induced by training, while preserving the use of an artificial dynamic cognition approach. However, evolutionary approaches are confronted with what [4] call the *hard problem of the enactive artificial intelligence* because it is necessary to associate phylogenetic mechanisms with the clarification of ontogenetic principles.

Before starting our explanation, we must mention that our work is similar to [7] who studies associative learning in evolved CTRNN. However, it is different because we use guidance to initiate the learning instead of discrete reward. Indeed, guidance is an evolutive mean to interact with the agent and to shape progressively its behavior. Here, we can use the metaphor of a child learning to ride a bike while being guided by an adult. If the child keeps his equilibrium, the adult stops to guide him progressively. Inversely, if the child loses his balance, the adult holds him. Progressively, the child will learn a new ability.

The section 2.1 presents our base experiment of learning by guidance. Models are then described in section 2.2 and section 2.3 presents the genetic algorithm used to set the parameters. The results are presented in the section 2.4.

2 Experiment

2.1 Principle

Let us consider an entity equipped with sensors functionally comparable to "eyes and ears". Its "ear" detects signals that make it change its orientation (right/left). Each "eye" detects the presence of one kind of light (A) or (B), but it initially does not display any particular behavior in presence of these lights. The signal sent to the ear will act as a guidance, so that an association between the signal received by the eyes and the one received by the ears is carried out dynamically in the interaction. In order to do so, we associate the presence of a light with a guidance moderated according to the effective behavior (for example: send signal turn to the right when light (A) is present and the entity is not turning to the right). The goal is that this guidance can attenuate gradually and that in the long term will no longer be necessary.

We study the behavioral enrichment at the individual level, i.e. at the ontogenetic level. Our artificial agents are equipped with a CTRNN. They have an effector to turn in both directions and three sensors that respectively detect a signal which corresponds to our guidance, the A light and the B light. The network is fully-connected. The agents must react to the guidance signal by turning to the right if the signal is negative and to the left if it is positive. We associate in an arbitrary way a light (A/B) with a side (left/right). The experiment consists in alternatively presenting the lights while guiding the agents according to the chosen association. The guidance signal is only present when the agent does not turn to the side associated with the light presented. However, a delay is introduced into the guidance mechanism to enable the dynamics of the system to take into account changes of perception. Thus, the guidance signal is not only delayed but also variable: it increases gradually if the agent behavior does not change. This adaptability leads us to speak about interactive guidance.

We seek to highlight the acquisition of a new behavior. This implies that the light presented does not condition *a priori* the behavior of the agents and thus that guidance must be necessary at the beginning of the experiment. If a new behavior conditioned by the light appears, then guidance became useless at the end of the experiment.

2.2 Model

The sensors mentioned previously and engines are coupled to the CTRNN. From then on, the term sensitive neuron will be used to designate a neuron which has an input coming from a sensor and the term motor neuron will be used to designate a neuron whose output is used by an effector. The cell potential y_i is governed by equation 1, where τ_i is the decay constant (range [0.5,2.1]), $w_{i,j}$ is the weight of synaptic connection from node i to node j (range [-8,8]), and I_i is an input from a sensor for a sensitive neuron. The firing rate z_i is calculated by using the equation 2, where b_i is the bias of the node i (range [-1,1]). The effector activation is computed by mapping the firing rate to the interval [-1,1] and multiplying by a gain (range [0,40]). In the same way, the input of the sensitive nodes is computed by multiplying the sensed value by a gain (range [0,40]). The sensed values are 1 or 0 on the "eyes" sensors depending on the presence of lights (A) and (B). For the "ear", the sensed value is negative for signal "turn to the left", positive for signal "turn to the right" and 0 for no signal.

$$\tau_i \frac{\delta y_i}{\delta t} = -y_i + \sum_j w_{j,i} z_j + I_i \tag{1}$$

$$z_i = \frac{1}{1 + e^{-(y_i + b_i)}} \tag{2}$$

$$\delta w_{i,j} = \eta(\alpha z_i z_j + \beta z_j + \gamma z_i) \tag{3}$$

The network plasticity, inspired from [15], corresponds to an hebbian rule (see equation 3), where η, α, β, γ are parameters (all in range [-1,1]) for each connection.

2.3 Parameters Setting

The parameters are: gains associated with the sensors and effector; for each neuron i, the decay constant τ_i and the bias b_i; for each connection, the parameters η, α, β and γ. They are determined by a genetic algorithm. The criteria to optimize by the algorithm is not a specific task but an ontogenetic development. The fitness used by the genetic algorithm is computed in 3 independent phases giving 3 scores : f_1, f_2 and f_3. In short, the three phases are:

1. Guidance reaction checking. During a trial, the entity is guided towards a side. The score of an entity for a trial is the time it has turned to the signaled side. This phase is made up of 20 trials alternating guidance towards the right and towards the left. Score f_1 is computed using

$$f_1 = \min_{i \in [0.20]} \left\{ val_i | val_i = \frac{\int_{ts_i}^{te_i} \Delta_{Turn}(t)dt}{te_i - ts_i} \left(1 - \frac{\int_{ts_i}^{te_i} \Delta_{Energy}(t)dt}{te_i - ts_i} \right) \right\}$$

where val_i is the score of the trial i which starts at ts_i and stops at te_i. Δ_{Turn} is the fitness component corresponding to the correlation between the guidance signal and the motor activity given by the equation

$$\Delta_{Turn}(t) = \begin{cases} 1 - \parallel M(t) - S(t) \parallel & : \quad M(t)S(t) > 0 \\ 0 & : \quad else \end{cases}$$

where $M(t)$ is the activity of the motor neuron mapped to the range [-1,1] and $S(t)$ is the guidance signal. Δ_{Energy} is used to avoid oscillating behaviors.

$$\Delta_{Energy}(t) = \begin{cases} 0 & : \quad M(t-1)M(t) > 0 \\ 1 & : \quad else \end{cases}$$

2. Guidance according to an association. Lights (A/B) are arbitrarily associated to sides (left/right). The duration of each light presentation is arbitrarily chosen in the range [1.5,4.5]. During a presentation, only one light is in the environment of the entity. After a time during which the behavior of the entity can stabilize (10 lights presentations), the guidance starts according to the association previously decided. This second phase is made up of 50 light presentations. For each presentation, the type of light is randomly chosen. Score f_2 is computed using

$$\begin{aligned} f_2 = \quad & (\min\{Score_A, Score_B\}) \\ & \min_{i \in [0.50]} \left\{ val_i | val_i = \frac{\int_{ts_i}^{te_i} \Delta_{Turn2}(t)dt}{te_i - ts_i} \left(1 - \frac{\int_{ts_i}^{te_i} \Delta_{Energy}(t)dt}{te_i - ts_i} \right) \right\} \end{aligned}$$

Δ_{Turn2} is given by

$$\Delta_{Turn2}(t) = \begin{cases} 1 & : \quad M(t)C(t) > 0 \\ 0 & : \quad else \end{cases}$$

where $C(t)$ stands for the association ($C(t) = 1$ when the presented light at time t is associated with side left and $C(t) = -1$ else). $Score_A$ and $Score_B$ are defined with

$$Score_X = \frac{Es_X - Ee_X}{Es_X}, X \in \{A, B\}$$

$Score_X$ measures the progress of the agent in associating the previously chosen side to light X. Es_X and Ee_X are respectively the mean value of the score at the beginning of the experiment and the mean value at the end.

$$Es_X = \frac{\sum_{i=10}^{20} \frac{1}{te_i - ts_i} \int_{ts_i}^{te_i} {\Delta_{Light}}^X(t)dt}{\sum_{i=10}^{20} X(t)}, X \in \{A, B\}$$

$$Ee_X = \frac{\sum_{i=40}^{50} \frac{1}{te_i - ts_i} \int_{ts_i}^{te_i} {\Delta_{Light}}^X(t)dt}{\sum_{i=40}^{50} X(t)}, X \in \{A, B\}$$

where $X(t)$ is 1 when light X is present, 0 else, and ${\Delta_{Light}}^X(t)$ is

$${\Delta_{Light}}^X(t) = \begin{cases} S(t) & : \quad C(t)M(t) > 0 \quad and \quad X(t) = 1, X \in \{A, B\} \\ 0 & : \quad else \end{cases}$$

3. Guidance according to the inverse association. The association used in phase 2 is inverted, then the same process is used to get a score f_3. By carrying out this inversion, we want to make sure the score of the agent depends on guidance associated with the light and not on a predisposition of the agent.

Between each phase, the neural network is reset. The final score of an agent is given by equation 4. A "good" agent is an agent which one can immediately guide and which is able to take into account a guidance to progressively associate a light with a side. The algorithm preserves the best individuals, in the proportion of a third. The second third is obtained by cloning the best individuals, each one of these clones systematically undergoing a mutation. All the parameters have the same mutation rate. The mutation consists in varying very slightly one of the parameters selected in a random way. The population is supplemented with a third of new individuals.

$$f = a * f_1 + b * min(f_2, f_3) \tag{4}$$

2.4 Results

The results presented here were obtained by applying the genetic algorithm to populations of 50 individuals, controlled by 6 neurons networks, during 30,000 generations. Each agent has 208 parameters which are directly encoded into genes (real values). Figure 1 illustrates how the best agent found by the genetic algorithm reacts to a guidance scenario when no light is in the environment. One guidance pulse is always sufficient to change the motors direction, however it may be necessary to repeat this guidance to stabilize the motor on the expected side. Figure 2 illustrates the protocol we used to observe the performance of the

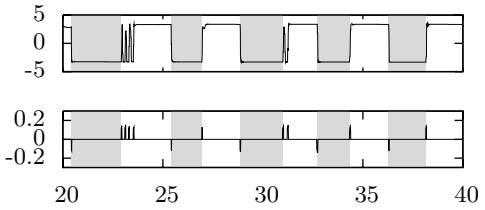

Fig. 1. Motor activity according to a guidance scenario. The first graph plots the motor activity and the second graph plots the guidance signal. Grey areas correspond to guidance towards the right whereas white areas correspond to guidance towards the left. The periods lasts are randomly chosen.

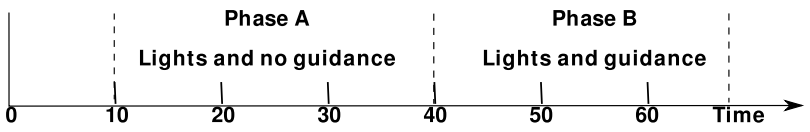

Fig. 2. Protocol used to observe the agents.

evolved agents. The experimenter decides on an association light-side. Phase A starts at time 10 in order to allow the network to reach a stationary mode. Lights are presented alternatively and there is no guidance. It allows us to check the non correlation between the motor activity and the lights before the training. Phase B, starting at time 40, corresponds to the training according to the chosen association. Lights are presented to the agent. The experimenter guides it if it is necessary. If the agent progresses, the experimenter intervenes less and less.

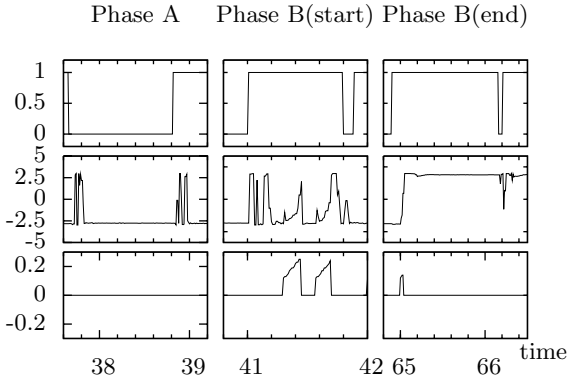

Fig. 3. Different time periods of one simulation show the ontogenetic evolution of the motor activity according to lights presentation.

Figure 3 shows the activity of the best evolved agent interacting with a virtual experimenter using the protocol described by figure 2. The association chosen is light A/left - light B/right. From top to bottom, the first graph corresponds to light presentations: 0 for light A, 1 for light B. A positive activity on the motor neuron (second graph) means that the agent is turning to the left, a negative activity that it is turning to the right. The third graph plots the guidance. At the end of phase A (left part on Figure 3), changes of lights induce a change in the motor activity. However, there is no direct correlation between the motor and the lights even if the agent exhibits a tendency to turn to the right. At the beginning of phase B (central part of Figure 3), the guidance causes the motor to turn towards the desired side, but it is necessary to repeat the guidance. At the end of phase B (right part of Figure 3), there is no need to guide the agent anymore as the motor activity is correlated to the lights (Obviously, changing side of the motor takes a little time, during which some little oscillations occur).

Figure 4 plots the motor activity compared to guidance during short time periods taken on the whole episode (not all time periods for the sake of clarity),

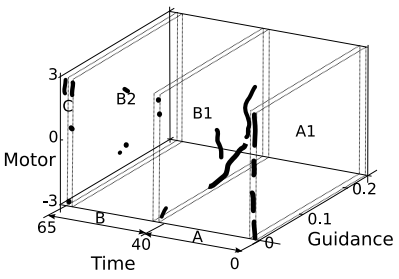

Fig. 4. Motor activity compared to guidance during some short time periods.

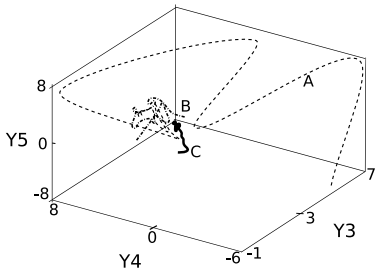

Fig. 5. Activity of the internal neurons (Y_4 and Y_5) compared to activity of the guidance sensitive neuron (Y_3).

but only when light A is in the environment. Light A has to be associated to side left, that means a positive motor activity. Set A1 corresponds to a short interval extracted from phase A, *i.e.* before the training. During this period, the motor activity is distributed between negative and positive values. Set B1 corresponds to the beginning of the training phase. At this time, motor activity is more negative with a high need of guidance. Set B2 is extracted from the end of the training. Motor activity is concentrated mostly on positive values. Set C corresponds to the first presentation of light A after the training period. During the considered time, motor activity is positive. Figure 5 illustrates how neural dynamics is shaped by training. Dynamics of neural outputs y_3, y_4, y_5 are represented at stages of presentation of light A. Y_3 is the output of neuron $n3$ which owns the guidance signal as input. It is surprising to observe that during phase A, this output varies widely. During phase B (guidance), trajectory joins a smaller space. Finally, this space is approximately kept during phase C (end of learning). This dynamic can be interpreted as the fact that during phase A, the system is very *receptive* to stimuli coming from guidance. Oscillations allow to explore a lot of states of the system. At the opposite, during guidance, the system falls into another lesser extended attractor, corresponding to an association side/light. However, this attraction depends on the guidance signal which depends itself on an arbitrary choice of the association used during the experiment. Similar curves are obtained for the presentation of light B.

3 Prospects

This work has focused on the evolution of the dynamics of a CTRNN during an associative learning task involving guidance. We have shown how an external influence, *i.e.* the guidance, may impact the internal dynamics. From a dynamical approach of artificial intelligence, it addresses the problem of sensory-motor acquisition of a dynamic system at an ontogenetic scale. To follow an enactive-like perspective, a long term goal is to reach co-evolutive mechanisms. This work is a

contribution in this direction, as it addresses the question of influencing the drift of a dynamic system through a external signal. Obviously, the task presented here is voluntarily very simple because we are more interested in the training than in the learned behavior itself.

To treat more complex tasks, it will be necessary to increase the sensory-motor capacities of the entity. For example, the design of an entity should require consideration of the entity's shape [11], the evolvability of the morphology of the neural network [6] and the imagination of slight variations of the complexity of the task. Moreover, the guidance strategy to adopt in order to learn a sensory-motor dynamics to an artificial entity is far from trivial. To address this problem, our perspective is to co-evolve guider and guided agents.

References

1. R.D. Beer. Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, 4(3):91–99, 2000.
2. R.A. Brooks. Intelligence without representation. *Artificial Intelligence*, 47:139–159, 1991.
3. E.A. Di Paolo. Homeostatic adaptation to inversion in the visual field and other sensorimotor disruptions. In J.A. Meyer, A. Berthoz, D. Floreano, H.L. Roitblat, and S.W. Wilson, editors, *From Animals to Animats 6. Proceedings of the VI International Conference on Simulation of Adaptove Behavior*, pages 440–449, 2000.
4. Tom Froese and Tom Ziemke. Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artif. Intell.*, 173(3-4):466–500, 2009.
5. K. Funahashi and N.i Nakamura. Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, 6:801–806, 1993.
6. F. Gruau. Automatic definition of modular neural networks. *Adaptive Behavior*, 3:151–183, 1994.
7. Eduardo Izquierdo, Inman Harvey, and Randall D. Beer. Associative learning on a continuum in evolved dynamical neural networks. *Adaptive Behavior*, 16(6):361–384, 2008.
8. B. McMullin. Thirty years of computational autopoiesis: A review. *Artificial Life*, 10:277–295, 2004.
9. Alberto Montebelli, Carlos Herrera, and Tom Ziemke. On cognition as dynamical coupling: An analysis of behavioral attractor dynamics. *Adaptive Behavior*, 16:182–195, 2008.
10. Mario Negrello and Frank Pasemann. Attractor landscapes and active tracking: The neurodynamics of embodied action. *Adaptive Behavior*, 16:196–216, 2008.
11. R. Pfeifer and G. Gomez. Interacting with the real world: design principles for intelligent systems. *Artificial life and Robotics*, 9(1):1–6, 2005.
12. Rolf Pfeifer and Josh C. Bongard. *How the Body Shapes the Way We Think, A New View of Intelligence*. MIT Press, 2006.
13. Jorge Simao. *Recurrent Neural Networks*, chapter Aperiodic (Chaotic) Behavior in RNN with Homeostasis as a Source of Behavior Novelty: Theory and Applications, pages 400–420. 2008.
14. F.J. Varela, E. Thompson, and E. Rosch. *The Embodied Mind*. MIT Press, 1993.
15. R. Wood and E. A. Di Paolo. New models for old questions: Evolutionary robotics and the 'a not b' error. In Springer-Verlag., editor, *Proceedings of the 9th European Conference on Artificial life ECAL 2007.*, 2007.