ESANN'2006 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2006, d-side publi., ISBN 2-930307-06-4.

Extended model of conditioned learning within latent inhibition

Nicolas Gomond and Jean-Marc Salotti

Laboratoire de Sciences Cognitives
Bordeaux Victor Segalen University - France

Abstract. Due to the various and dynamic nature of stimuli, decisions of intelligent agents must rely on the coordination of complex cognitive systems. This paper precisely focusses on a general learning architecture for autonomous agents. It is based on a neural network model that enables the specific behaviours of classical conditioning and a biologically inspired attentional phenomenon called latent inhibition. We propose a neural network implementation of an extended model of classical conditioning and present some results.

1 Introduction

Our objective is to design a global learning architecture exploiting the connectionist techniques such as artificial neural networks. It should integrate temporal constraints and exhibit some sensorimotor learning capacities. Our main goal is in fact twofold: first, we would like to validate a neurobiological hypothesis of associative learning and latent inhibition. On the second hand, we are interested in a cognitive learning model that can be implemented to command robots.

One of the most studied elementary cognitive processes to understand animal behaviours is classical conditioning. These mechanisms are dependent on the obtained rewards. A latent inhibition experiment is often performed to study the neurobiological system involved in this learning process. Latent Inhibition (LI) is the phenomenon of reduction in the conditioning capacity of a neutral stimulus if it were initially presented alone several times without reinforcement [1]. In other words, preexposure (PE) of the to-be-conditioned stimulus before the conditioning phase induces some delay for the association learning. For example, LI can easily be observed when an unexpected indigestion occurs after a traditional meal. This natural phenomenon is considered to play an important behavioural role because it allows exclusion of useless information from the consciousness, avoiding the learning of irrelevant stimuli and thus improving the action selection mechanism. Schmajuk et al.[2, 3] propose a dynamic system of classical conditioning, which simulates certain behavioural effects of LI. Their model (SLG) assumes that the association effectiveness of a conditioned stimulus (CS) with an unconditioned one (US) is proportional to the novelty of the CS. The strength of the internal representation of the CS decreases during preexposure phase proportionnally to its novelty. We implemented our architecture using the concept of the SLG model. While increasing the similarities with the biological networks, our model also takes into account timing effects ; various and numerous input stimuli can be used during an extended period of time and

many situations of conditioning tests can be simulated compared to the SLG model or other classical conditioning models [4, 5, 6]. It is indeed interesting to test the global learning architecture within a larger behavioural framework in a real environment. Thus, our architecture has been designed for the improvement of the action-selection mechanism, for the processing of the stimuli and also for the perception-reward-action learning.
The global learning architecture is presented in section 2.1 and some results are presented and discussed in section 3.

2 Model

2.1 Main features

We propose a learning architecture that maintains an agent in a functional state according to internal variables (needs or motivation) and to the external environment stimuli. Our architecture is inspired by several computational models [7, 8].

2.2 Architecture

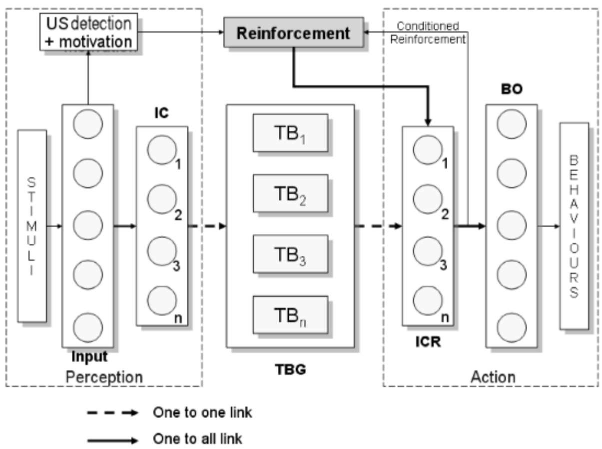

Fig. 1: Global learning architecture. US : Unconditioned Stimulus, IC : Input Classificator, TBG : Time Base Group, TB : Time Battery, ICR : Internal Conditioned Response, BO : Behavioural Output

There are three distinct modules in our model (see Fig. 1), the first one for the perception of stimuli, the second one for associative learning and the third for the selection of behaviours. The first layer of the perception part can be considered as the result of the basic preprocessing phase of the input stimuli. Indeed on this level, inputs of our network contain all information concerning stimuli (intensity

and reinforcing character for example). The perception module is in charge of permanent and unsupervised learning with incremental memorization and detection of novelty. An ART-1 neural network [9] has been chosen for the implementation of this module. It is able to adapt to non familiar inputs by creating new categories (plasticity). And it can also adapt the classes already learned while degrading already memorized information (stability).
After this classification, each output neuron of the IC layer (Input Classificator) is used as input to a single "Time Battery" unit (TB) in the "Time Base Group" (TBG). A TB acts as delay neurons endowed with different time constants. Such delay units have been proposed by several authors to obtain different timing properties [10, 11, 12]. A TB performs a spectral decomposition of the signal. An input to a specific battery implies both a reset of any previous activity in this battery and an initialization of the spectral timing activity. The main interest of this set of spectral neurons is to allow a compact and robust coding of time. The first neuron of the battery has a very strong but short activity, which very quickly reaches its maximum. The following neurons have decreasing activities which reach their maximum later but last longer. There are ten cells per each battery. The activation law of a given neuron j follows a Gaussian distribution (Eq. 1).

$$Act_j(t) = \frac{1}{m_j} \cdot \exp^{-\frac{(t-m_j)^2}{2 \cdot \sigma_j}} \tag{1}$$

t is the time in second, m_j is the time constant for maximum intensity and σ_j is the standard deviation.
Globally, TBG is the key element of the conditioning, the prediction and the learning between perception, motivation and action. The output signal of a battery is represented by an integrator neuron. It can be interpreted as the Internal Conditioned Response (ICR) corresponding to the specific input stimulus. That response generates two different activities. Firstly, it is a conditioned reinforcement signal for another stimulus, which is about to come later through another TB. And in another side, it is the input signal of a neural network, which is dedicated to the selection of action. This layer represents the behavioural part of the system (BO : Behavioural Output). It determines the final behaviour of the agent.

2.3 Learning and predictive system

Another important aspect of our architecture is that a reinforcement neuron has been added. It positively acts on each output neuron of all TBs. It is activated when a reward signal is received. The source of the reward can be external by means of an unconditioned stimulus, internal in the case of a motivation signal or if an already conditioned reinforcement is provided by ICR neurons. When the reinforcement neuron is activated, the weights of the connection between the conditioned response neuron (ICR) and the most active TB neuron at that time are updated according to Eq.2:

$$W_t = W_{t-1} + \triangle \tag{2}$$

with $\triangle$, the weight variation (Gaussian distribution):

$$\triangle = k \cdot \exp^{-\alpha \cdot (W_{t-1} - \mu)^2} - \rho \tag{3}$$

$$with \begin{cases} \rho = +0.01, & \text{for } W_t > 0; \\ \rho = -0.01, & \text{otherwise.} \end{cases}$$

α and μ are scale and location parameters, respectively. ρ is a leak.

The other weights of the TB are decreased according to Eq.4. On the other hand, if a stimulus is not activated when there is a reset of activity of the corresponding TB, all weights are decreased according to Eq.4 :

$$W_t = W_{t-1} - \triangle \tag{4}$$

Furthermore, we propose additional mechanisms to take into account the features of LI and other important properties of classical conditioning.

- extinction : synaptic weights progressively return to their basal state if they are not reactivated.
- conditioned reinforcement : a sufficient answer can be used as a reinforcement signal for another stimulus. Thus, it is possible to learn a sequence of stimuli.
- selection of action : coactivation of one unconditioned response and one conditioned response involves a modification of the weights into the BO network. This sensorimotor coupling allows the learning of the adequate behaviour when the conditioned stimulus is presented alone.

3 Results and discussion

Our network has been tested in multiple cases of conditioning tests. For example, a first neutral stimulus (CS) is presented in input layer. Then there is activation of the neurons of the corresponding TB. After a delay, a second stimulus (reinforcing = US) is presented. We repeat the same sequence several times. The test phase consists in presenting the CS alone and observing the response of the corresponding ICR neuron. Figure 2 shows the internal conditioned responses obtained after twenty associations between CS and US.
The second simulation is a classical conditioning test with the latent inhibition phenomenon. A first stimulus is presented alone several times with random delays. The same sequence operated in the first simulation is then performed and the test of conditioning finally occurs. Figure 3 shows the results of the same experiment except that CS-US associations are preceded by preexposure of ten CS. As can be observed, the internal conditioned responses are delayed. The effect of latent inhibition is appropriately noted after successive presentation of the CS alone. The repeated presentation of the CS alone induces a progressive decrease of the response until the loss of conditioning. The sequence CS-delay-US produces an increase of the weight of the TB neuron corresponding to the

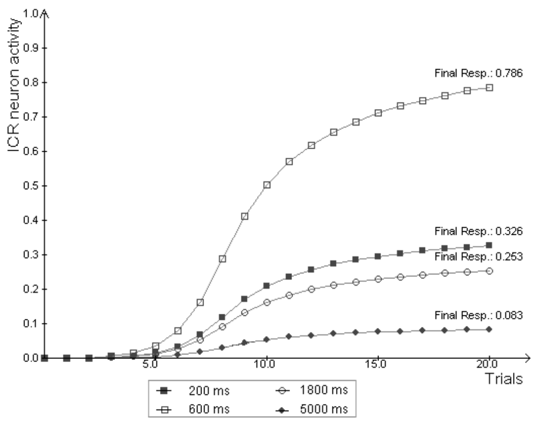

Fig. 2: Classical conditioning test : Conditioned response after 20 CS-US associations with different CS-US delays (200, 600, 1800 and 5000 ms)

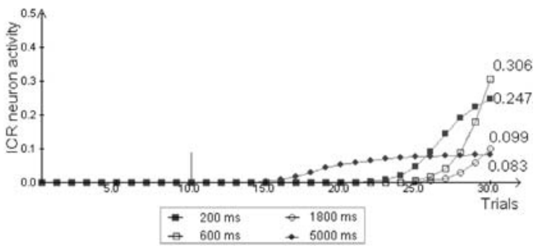

Fig. 3: Latent Inhibition : Conditioned response after 20 CS-US associations preceded by 10 CS with different CS-US delays (200, 600, 1800 and 5000 ms)

delay. Thus, gradually, the resulting activity becomes sufficient to release a response in the conditioned response neuron. Our results are in adequacy with the behavioural responses obtained with animals [13]. In fact, we observed several characteristics of classical conditioning and latent inhibition in many behavioural experiments.

- The larger the delay between the CS and the US, the longer it takes to obtain a conditioning and the weaker the internal conditioned response. This is due to the characteristics of TB neurons. An important delay between the CS and the US involves a modification of the weight of the TB neuron with a long but weak activity.
- During the test phase, we obtain a response of the conditioned response neuron exactly when the theoretical presentation of the US is expected.
- Many other experiments have been carried out with our model with various

ESANN'2006 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2006, d-side publi., ISBN 2-930307-06-4.

delays, reinforcements, stimuli and experimental situations (blocking test, CS-CS or CS-US associations, context tests. . .).

4 Conclusions and perspectives

From behavioural experiments on animals and also thanks to various existing models we defined an artificial neural network architecture allowing the classical conditioning, the learning of the delay and the sequence of stimuli. Our model embeds certain interesting features. Firstly, it is a real time model that adaptively builds the relationships between stimuli and reinforcements. Furthermore, it enables the effect of latent inhibition, which is an important characteristic in the conditioning and action selection processes. Moreover, it provides a suitable framework for the learning of specific and appropriate behaviours in complex environnements. Thanks to the first encouraging results, we will soon realize experiments using a mobile robotics simulation software. Indeed, our model will be implemented on Webots (c) software [14]. Then, we will be able to study the behaviours of an autonomous robot using our conditioning learning model and a dedicated action module.

References

[1] R.E. Lubow. Latent inhibition. *Psychol. Bull.*, 79:398–407, 1973.

[2] N.A. Schmajuk, Y. Lam, and J. A. Gray. Latent Inhibition: a neural network approach. *Journal of experimental psychology: animal behavior processes*, 22(3):321–349, july 1996.

[3] N.A. Schmajuk, C.V. Buhusi, and J. A. Gray. Psychopharmacology of latent inhibition: a neural network approach. *Behavioural Pharmacology*, 9(8):711–730, 1998.

[4] R.A. Rescorla and A.R. Wagner. *A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement*, volume Classical conditioning II: current research and theory. New York: Appleton-Century-Crofts, 1972.

[5] Sutton R. and Barto A., editors. *A temporal difference model of classical conditioning*, Proceedings of the ninth conference of the cognitive science society, 1987.

[6] P.F.M.J. Verschure and A.C.C. Coolen. Adaptive fields: distributed representations of classically conditioned associations. *Network*, 2:189–206, 1991.

[7] P. Gaussier and S. Zrehen. *A constructivist approach for autonomous agents*. Willey and Sons, n. thalman edition, 1994.

[8] B. Girard, V. Cuzin, A. Guillot, K.N. Gurney, and T.J. Prescott. Comparing a brain-inspired robot action selection mechanism mith winner-takes-all. In B. Hallam, D. Floreano, J. Hallam, G. Hayes, and J.-A. Meyer, editors, *From Animals to Animats 7*. The MIT Press, 2002.

[9] S. Grossberg. Competitive learning : from interactive activation to adaptive resonance. *Cognitive science*, 11:23–63, 1987.

[10] S. Grossberg and J.W.L. Merrill. A neural model of adaptively timed reinforcement learning and hippocampal dynamics. *Cognitive Brain Research*, 1:3–38, 1992.

[11] J.P Banquet, J.L. Contreras-Vidal, P. Gaussier, and Y. Burnod. *Fundamentals of neural network modelling for neuropsychologists, chapter The cortical-hippocampal system as a multirange temporal processor : A neural model*. MIT Press, 1996.

[12] D. Bullock, J.C. Fiala, and S. Grossberg. A neural model of timed response learning in the cerebellum. *Neural Networks*, 7:1101–1114, 1994.

[13] R.E. Lubow. *Latent inhibition and conditioned attention theory*. Cambridge, 1989.

[14] Heuding J.-C., editor. *Webots: symbiosis between virtual and real mobile robots*, Proceedings of the First Int. Conference on Virtual Worlds, 1998.