

# Emergence of Rules in Cell Assemblies of fLIF Neurons

Roman V. Belavkin and Christian R. Huyck<sup>1</sup>

**Abstract.** Inspired by biological cognition, CABOT project explores the ways symbolic processing can emerge in a system of neural cell assemblies (CAs). Here we show how a stochastic meta-control process can regulate learning of associations between the CAs, the neural basis of symbols. An experiment illustrates the learning between CAs representing conditions actions pairs, which leads to CA-based representations of ‘if-then’ rules.

## 1 INTRODUCTION

Previously, the authors have demonstrated how states in cell assembly (CA) neural system can be controlled and used to perform a typical symbolic task (counting) [5]. This work has developed into a much more ambitious project called CABOT, where the same principles are applied in a system, based entirely on CAs, that integrates elements of vision, categorisation, natural language processing and learning in a virtual environment. This paper presents a part of this project — learning the connections between different CAs — that combines symbolic representations into logical rules.

## 2 OVERVIEW OF THE ARCHITECTURE

Our system uses fatiguing, leaky, integrate and fire (fLIF) neurons [4], an extension of LIF neurons [6]:

**Integrate and fire** — the neuron ‘fires’ if its action potential,  $A$ , exceeds threshold  $\theta$ , where  $A = (w, x) = \sum_{i=1}^k w_i x_i$  (integrator),  $w, x \in \mathbb{R}^k$  are the weights and the stimuli vectors. The weights  $w_i$  adapt according to the compensatory learning rule [4], which is an implementation of Hebbian learning [3].

**Leak** and accumulation of potential,  $A_{t+1} = \frac{A_t}{d_t} + (w_t, x_t)$ , where  $d_t = \infty$  if fired at  $t$ ;  $d \geq 1$  otherwise.

**Fatigue** makes the threshold dynamic,  $\theta_{t+1} = \theta_t + F_t$ , where  $F_t = F_+ \geq 0$  if fired (*fatigue*);  $F_t = F_- < 0$  otherwise (*recovery*).

**Cell assemblies** are reverberating groups of neurons [3], and they are believed to be the neural basis of symbols in human mind. Our system is based on networks of sparsely connected neurons. The topology of the networks is pre-defined by some random pattern, and it can be highly recurrent. When enough neurons fire to start the reverberating circuit, the CA ignites, and its *persistence* is an important property of CAs’ dynamics. The fatigue and recovery rate parameters affect the persistence. A CA can be extinguished by another CA, which can ignite due to the change of the external pattern.

A network with several CAs encoding a set of external patterns is referred to as a *module*. Several modules can be interconnected to create more complex systems. For example, a system of 7 modules and 40 CAs was used to implement a simple counting task [5]. More

complex systems have been used to parse natural language and implement finite state automata. The next stage in the development of the project is the ability to learn the connections between different modules, the focus of this paper.

## 3 STOCHASTIC META-CONTROL

Although the connections between the correlated cells are strengthened via Hebbian learning, it is the meta-process that controls which neurons fire and thus which connections are supported. The meta-process is based on stochastic control of action-selection algorithms, implemented earlier by the authors in cognitive architectures [1] and which are based on the following result of information theory.

Given utility function  $u : \Omega \rightarrow \mathbb{R}$ , the goal is to find probability distribution,  $p$  on  $\Omega$ , that maximises the expected utility  $E_p\{u\} = \int p(u) u$  under additional constraints. This distribution is

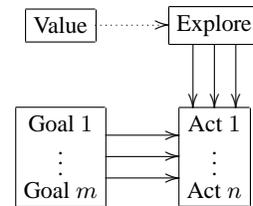
$$p(\omega) = q(\omega) e^{\beta u(\omega) - \Gamma(\beta)} \quad (1)$$

where  $q(\omega)$  is the reference (prior) distribution,  $\Gamma(\beta) = \ln \int_{\Omega} q(\omega) e^{\beta u(\omega)}$ , and  $\beta$  is the Lagrange multiplier, defined from constraints on information ( $I(p, q) \leq I < \infty$ ) or on the expected utility ( $E_p\{u\} \geq U > -\infty$ ):

$$\beta(U) = \frac{dI(U)}{dU}, \quad I(U) = \sup_{\beta} [U\beta - \Gamma(\beta)] \quad (2)$$

Function  $\beta(U)$  is strictly increasing, and for  $\beta > 0$  the optimal distribution (1) has non-zero values ( $p(\omega) > 0$ ) for all  $\omega \in \Omega$  such that  $u(\omega) > -\infty$ . Thus, the optimal distribution describes stochastic process, where all  $\omega$  are randomised by the control parameter  $\beta > 0$ , or its inverse  $T = \beta^{-1}$  called the *temperature*.

**Value-Explore Topology.** Problems of optimal control often involve maximisation of utility over set  $\Omega = X \times Y$ , where  $X$  is the set of observations (e.g. goals), and  $Y$  is the set of controls (e.g. actions). In our system, these sets are represented by two modules, Goals and Actions, where CAs represent conditions and actions respectively. Thus,  $\omega \in \Omega$  are condition-action pairs  $(x, y) \in X \times Y$ .



Initially, the modules are set up with excitatory connections from every  $x \in X$  to all  $y \in Y$ . Thus, given some goal, any action can be triggered. Due to the Hebbian learning, the connections  $x \rightarrow y$  between CAs that have fired together are reinforced, giving the pair a

<sup>1</sup> Middlesex University, London NW4 4BT, UK

higher chance to ignite in the future. Thus, due to Hebbian learning, the system can learn some random relation  $R \subset X \times Y$  (set of rules), which may not be optimal. Learning of only a particular (optimal) relation is supported by the meta-process that involves two additional modules: Value and Explore.

The activity of the Value module represents the values of utility (higher activity corresponds to higher utility). The average activity of the module corresponds to constraint  $U$  in equation (2). The input of the module can be configured according to the application. For example, it may receive inputs from the sensory system representing agent's preference on the states of the environment.

The purpose of the Explore module is to randomise the activity of the Action module. Cells in this module are spontaneously firing, and the module sends excitatory connections to all CAs in the Action net. Thus, the Explore module can trigger randomly any Action CA, and this process has no memory. The module implements the effect of parameter  $\beta > 0$  in equation (1) (or the temperature  $T = \beta^{-1}$ ).

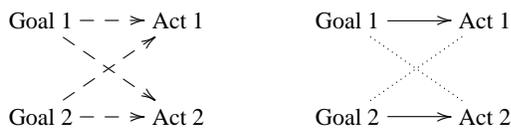
The Value module sends inhibitory connections to Explore, so that high activity of Value inhibits the activity of Explore. This implements the monotonic relation between constraint  $U$  and  $\beta$  in equation (2), and it allows for a very simple yet effective learning scheme. If a particular goal-action pair  $(x, y)$  results in a high utility, then the Value module inhibits Explore, and the  $(x, y)$  pair is allowed to persist longer. Since high utility pairs  $(x, y)$  on average co-fire longer than low utility pairs, their connections increase relative to others due to the compensatory Hebbian learning rule.

This way, the meta-process supports learning of the optimal relation  $R \subset X \times Y$ . As a result, the average activity of the Value module ( $U$ ) increases with time, while the activity of the Explore module ( $T = \beta^{-1}$ ) decreases. The system makes a transition from stochastic to an almost deterministic rule-based system.

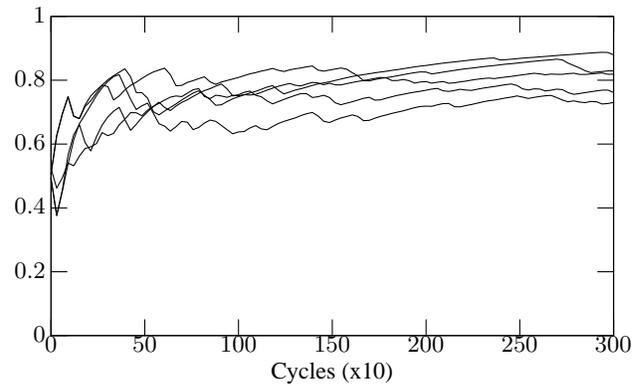
**The biological plausibility** of this topology is supported by studies of the reward path and tonically active cholinergic neurons in the basal ganglia and striatal complex [2]. These neurons account for a small proportion of the connections, and they are quite uniform and nontopographic. These neurons may play the role of stochastic noise, and their activation is reduced when the reward path is activated.

## 4 EXPERIMENT: LEARNING DICHOTOMIES

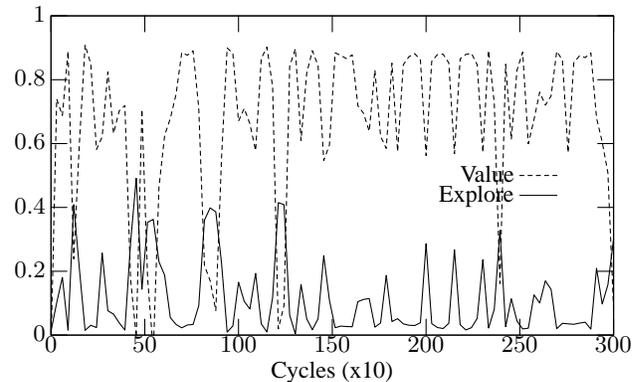
The code of the system and the experiment described is available at <http://www.cwa.mdx.ac.uk/CABot/CANT.html>. In this simple experiment, there are two CAs in the Goal and two CAs in the Action modules. Each module consisted of 800 cells, with 400 cells in each CA. The modules were set up with low weight excitatory connections from every goal CA to all action CAs, shown by dashed arrows on the left diagram below. The task was to learn two rules, shown by two solid arrows on the right diagram.



The training procedure consisted of a random presentation of an input pattern activating one of the goal CAs every 100 cycles. Figure 1 shows the proportion of the correct actions selected (ordinate) as a function of cycles (abscissa). The chart shows the results of five simulations. Initially the system makes only half of the choices correctly. After 3000 cycles, the proportion of correct choices increases to 70–90%. Figure 2 shows the percentage of neurons firing per cycle



**Figure 1.** The proportion of correct action choices (ordinate) as a function of cycles (abscissa). The curves represent results of different trials.



**Figure 2.** Activities of the Value and Explore modules in one experiment.

in the Value and the Explore modules in one of the experiments. As desired, an increase of the Value activity coincides with the decrease of the Explore.

The implementation of the meta-process for rule acquisition in our system is an important step in its evolution creating new opportunities and improving our understanding of biological cognition.

## ACKNOWLEDGEMENTS

This work was supported by EPSRC grant EP/DO59720.

## REFERENCES

- [1] R. V. Belavkin, 'Acting irrationally to improve performance in stochastic worlds', in *Proceedings of AI-2005, the 25th SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence*, eds., M. Bramer, F. Coenen, and T. Allen, pp. 305–316, Cambridge, (December 2005). Springer.
- [2] R. Granger, 'Engines of the brain: The computational instruction set of human cognition', *AI Magazine*, **27**(2), 15–32, (July 2006).
- [3] D. O. Hebb, *The Organization of Behavior*, John Wiley & Sons, New York, 1949.
- [4] C. Huyck, 'Hierarchical cell assemblies', *Connection Science*, (2007).
- [5] C. Huyck and R. V. Belavkin, 'Counting with neurons, rule application with nets of fatiguing leaky integrate and fire neurons', in *Proceedings of the Seventh International Conference on Cognitive Modeling*, eds., D. Fum, F. D. Missier, and A. Stocco, Trieste, Italy, (April 2006). Edizioni Goliardiche.
- [6] W. Maas and C. Bishop, *Pulsed Neural Networks*, MIT Press, 2001.