The Emergence of Rules in Cell Assemblies of fLIF Neurons

Roman V. Belavkin and **Christian R. Huyck** [1]

Abstract. There are many examples of intelligent and learning systems that are based either on the connectionist or the symbolic approach. Although the latter has been successfully combined with statistical learning to create a hybrid system, it is not clear how symbolic processing can emerge from a connectionist system. The human mind is a living proof that such a transition must be possible. Inspired by biological cognition, our project explores the ways symbolic processing can emerge in a system of neural cell assemblies (CAs). Here, we present a meta–process that regulates learning of associations between the CAs. The process is compared with stochastic learning theory, and its outcome is a set of optimal rules implemented in simulated neurons and learned by Hebbian adaptation of synaptic weights. A neural simulation shows rules can be learned.

1 INTRODUCTION

In recent decades, theories of cognition have been developed using different paradigms — some are based entirely on studies and simulations of biological neural processing, while others pursue a more abstract approach by simulating the behaviour. The former facilitated the solution of a great variety of engineering problems (e.g. signal processing, pattern recognition), while the latter have revolutionised cognitive psychology [17]. Despite the successes, the process of unification of neural and symbolic cognitive systems has been slow even though human cognition — the main subject of both approaches — is a clear indication that they are two sides of the same coin.

Although a single neuron can classify a large number of patterns, it is believed that groups of connected cells called *cell assemblies* (CAs) form the basis of human cognition [7, 18]. However, recent advances in modelling human–level cognition have mostly been made using symbolic cognitive architectures, such as SOAR [17] and ACT–R [1]. The success of the latter can be explained largely because it uses a hybrid approach, where symbols are applied selectively based on statistical associations and other sub–symbolic computations. The work described in this paper is part of a project attempting to achieve complex symbolic processing and learning in a connectionist system.

Previously, the authors have demonstrated how states in a CA–based system can be controlled and used to perform a typical symbolic task (counting) [11]. 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 virtual environments. This paper presents a part of this project — learning the connections between different CAs — that implement combinations of symbolic representations and ultimately the emergence of logical rules. In this system, learning is modification of synaptic weights via a Hebbian rule, but a meta–learning process emerges from the interaction of large groups of neurons in different modules.

In the next section, the model of a fatiguing, leaky, integrate and fire (fLIF) neuron, CAs, and the use of CAs as symbols are described. In the third section, information–theoretic analysis of stochastic learning will be outlined and its implementation in our system will be presented. The fourth section will present a simple experiment illustrating the working of the system and its relation to other works will be discussed. The biological plausibility of the learning process will also be considered in the last section.

[1] Middlesex University, London NW4 4BT, UK

2 OVERVIEW OF THE ARCHITECTURE

Below is an overview of the neural model and its parameters. A more detailed presentation can be found in [10].

2.1 Fatiguing Leaky Integrate and Fire neurons

Biological neurons are complex systems, and the levels of details varied significantly even in the early models [8, 16]. Our system uses spiking, fatiguing, leaky, integrate and fire (fLIF) neurons, an extension of the LIF model [15]. Our model is a compromise between computational efficiency and biological plausibility reflecting properties that are, in our opinion, important for the emerging dynamics.

The 'integrate and fire' component is based on the classical idea [16] that the neuron 'fires' if its action potential, A, exceeds a certain threshold value θ. The action potential is a function of the inner product (integrator) $(w, x) = \sum_{i=1}^{k} w_i\, x_i$, where $x \in \mathbb{R}^k$ is the stimulus vector (pre–synaptic), and $w \in \mathbb{R}^k$ is the synaptic weight vector of the neuron. Here, $\mathbb{R}^k$ is k–dimensional Euclidean space, where k is the number of synapses to the neuron. If x is the output of all pre–synaptic neurons, then x is a binary vector.

The action potential depends on the pre and post–synaptic activity over several time moments:

$$A_{t+1} = \frac{A_t}{d_t} + (w_t, x_t)\,, \quad d_t \equiv \begin{cases} +\infty & \text{if fired} \\ d \geq 1 & \text{otherwise} \end{cases}$$

Thus, the action potential is accumulated if the neuron does not fire. Parameter $d > 1$ allows for some of this activation to 'leak' away. This is the LIF model.

The threshold of a neuron is also dynamic

$$\theta_{t+1} = \theta_t + F_t\,, \quad F_t \equiv \begin{cases} F_+ \geq 0 & \text{if fired} \\ F_- < 0 & \text{otherwise} \end{cases}$$

where values F_+ and F_- represent the *fatigue* and fatigue *recovery* rates. Thus, if a neuron fires at time t, its threshold increases, and it is less likely to fire at time $t + 1$.

Finally, the weights w_t can adapt according to the compensatory learning rule [10], which is an implementation of the Hebbian principle [7], where w_{t+1} depends on the correlation between the pre–synaptic, x_t, and the post–synaptic, y_t, activities. One can see that the post–synaptic activity is a non–linear functional of the pre–synaptic activity: $y_t : x_t \to \mathbb{R}$.

2.2 Cell assemblies

The system is based on networks of sparsely connected cells. The topology of the network is pre–defined by some random pattern, and it can be highly recurrent, similar to the Hopfield networks [9]. Unlike the Hopfield nets, however, the links are unidirectional, making our model more biologically plausible.

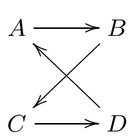
The non–linearity of the cells in the network leads to a complex dynamics similar to that in attractor nets with some of the states being more probable. These stable states can be characterised by groups of cells that remain significantly more active than the other cells in the system. Such reverberating groups of cells are often referred to according to Hebb [7] as *cell assemblies* (CAs). Hebb also felt that CAs were the neural basis of symbols.

An important property of CAs' dynamics is their persistence [13]. When enough neurons fire to start the reverberating circuit, the CA ignites. Once ignited, the activity within the cells in a CA may be sufficient to support itself. Many variables can contribute to this effect. In particular, the fatigue and recovery rate parameters in our system effect persistence. A CA can be extinguished by another CA, which can ignite, for instance, due to the change of the external pattern.

Sensory stimuli of fLIF neurons allows Hebbian learning to encode and store information about that stimuli. Note that CAs are not necessarily disjoint sets of cells. A single cell may be a member of several overlapping CAs. This feature can be used to encode hierarchies of patterns [10].

2.3 Symbolic processing with CAs

In the system described, 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. It was demonstrated earlier that state transitions in such systems are sufficiently controllable to implement a broad range of algorithms similar to symbolic systems. For example, a simple system with four CAs A, B, C and D oscillating in the $ABCD$ order can be created using two modules AC and BD, where the CAs are linked by excitatory connections as shown below.



The same principle can be used to simulate more complex behaviour. For example, a system of 7 modules and 40 CAs was used to implement a simple counting task [11]. More complex systems have been used to parse natural language and implement finite state automata.

Although the CAs within the individual modules of these systems could form via Hebbian learning between the cells in the network, the connections between the modules had to be set in a controlled way for the system to operate in the desired manner. The next stage in the development of the project is the ability to learn the connections between different modules, and that is the main focus of this paper. Before describing the process, we note that learning of the connections between different modules involves a meta–process. Indeed, 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. This meta–process is based on stochastic learning theory, which is briefly outlined in the next section.

3 STOCHASTIC LEARNING

The meta–process for learning the connectivity between the modules is based on the stochastic action–selection algorithms implemented earlier in cognitive architectures and stochastic symbolic systems [3, 4]. Theoretical foundations of this theory are based on the variational problems of information theory [19, 20], a generalisation of which is outlined below.

3.1 Optimisation with information constraints

Rational action selection is related to the theories of choice and optimisation. Fundamental in the theory of choice is the concept of a preference relation on a set (total and transitive binary relation). Often, the preference relation can be represented by a monotone function $u : \Omega \to \mathbb{R}$ referred to as the *utility*, and the choice problem is solved by maximisation of $u(\omega)$ (i.e. optimisation).

Under uncertainty, the choice problem is solved by using the preference relation on set P of *all* probability measures, which are non–negative functions $p : \mathcal{F} \to [0,1]$ defined on the σ–algebra $\mathcal{F}$ of Ω, and such that $p(\Omega) = 1$. The preference relation on P is induced by the *expected utility* $(p,u) = \int_\Omega u(\omega)p(\omega)d\omega$, so that for any p, $q \in P$, measure p is preferred if $(p,u) \geq (q,u)$, and it is the classical Bayesian estimation procedure [23, 24].

More generally, the problem of optimisation under uncertainty can be viewed as maximisation in the conjugate space. Indeed, given a Banach space U, the conjugate space V is the totality of all linear functionals (v,u), where $(\cdot,\cdot)$ is the inner product. Thus, given utility function $u \in U$, the maximisation of the expected utility corresponds to finding the maximum element $p \in P \subset V$, where P is the set of all probability measures.

It is often the case that the choice set under uncertainty is not the entire set P, but some subset of it defined by constraints. In particular, adaptive and learning problems are concerned with constraints on information, which can be defined in general form using the *information divergence*:

$$I(p,q) = \int_\Omega \ln \frac{dp}{dq}\, p(d\omega) \tag{1}$$

where measures p, $q \in P$ are such that p is absolutely continuous with respect to the reference measure q, and dp/dq is the Radon–Nikodym derivative. Note that for $q = \mathrm{const}$, information divergence corresponds to minus entropy, and when p and q are the conditional and the marginal probabilities respectively, then $I(p,q)$ is the Shannon information.

The important properties of information divergence is that it is convex, non–negative and its minimum is achieved for $p = q$ (see [14]). The maximum of $I(p,q)$, which can be infinite, is achieved for $p \to \delta_{\omega\omega'}$, which are the probability measures concentrated entirely on single elements of Ω (here $\delta_{\omega\omega'}$ is the Kronecker symbol). Thus, the constraints $I(\mu,\nu) \leq I = \mathrm{const} < \infty$ define some convex set $P' \subset P$, and the problem can be formulated as the following convex optimisation problem with information constraints:

$$\max_{p \in P'} (p,u)\,, \qquad P' \equiv \{p \in P : I(p,q) \leq I < \infty\}$$

This variational problem can be solved via the standard procedure of Lagrange multipliers. The solution is the probability measure:

$$p(d\omega) = q(d\omega)\, e^{\beta u(\omega) - \Gamma(\beta)} \tag{2}$$

where $\beta \geq 0$ is the Lagrange multiplier defined by $I(\mu, \nu) = I$, and $\Gamma(\beta) = \ln \int_\Omega e^{\beta u(\omega)} q(d\omega)$ due to the normalisation condition ($p(\Omega) = q(\Omega) = 1$). Note that the Gibbs distribution, known from thermodynamics, is a special case of function (2) (i.e. when $q(d\omega) = \text{const}$). Probability measure (2) corresponds to the maximum of the expected utility when the information divergence is bounded above $I(p, q) \leq I$. Furthermore, the problem of minimisation of information divergence with constraints on expected utility $(p, u) \geq U$ has the solution in exactly the same form, but parameter β determined from condition $(p, u) = U$. The relation between the information–utility constraints and parameter β, defining the optimal solution, can be expressed using the Legendre–Fenchel transform of potential $\Gamma(\beta)$:

$$I(U) = \sup_\beta [U\beta - \Gamma(\beta)]\,, \quad \Gamma(\beta) = \sup_U [\beta\, U - I(U)] \tag{3}$$

which corresponds to the following canonical equations

$$U(\beta) = \frac{\mathrm{d}\Gamma(\beta)}{\mathrm{d}\beta}\,, \quad \beta(U) = \frac{\mathrm{d}I(U)}{\mathrm{d}U} \tag{4}$$

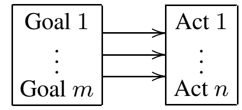
In particular, the second equation above suggest that an increase of the expected utility and information during learning corresponds to a positive value of parameter β. Moreover, $\Gamma(\beta)$ is convex, and therefore $I(U)$ is convex as well (property of the Legendre–Fenchel transform). Thus, $\beta(U)$ is a non–decreasing function. One can see from (2) that for all $d\omega \subseteq \Omega$ such that $q(d\omega) > 0$ and $u(\omega) > -\infty$, the condition $\beta > 0$ implies $p(d\omega) > 0$ as well. Thus, the optimal solution for optimisation with information constraints is a stochastic process (i.e. non–deterministic, or $p(d\omega) \neq 1$ for all $\omega \in \Omega$).

It has been known for a long time that stochastic algorithms outperform deterministic strategies in problems involving information constraints, such as the problems of rare event estimation and adaptive problems. The Gibbs distribution has been used in many optimisation techniques and machine learning algorithms to control exploration (e.g. simulated annealing). A similar random strategy has been employed by the ACT–R cognitive architecture [1] to simulate statistical learning of human subjects and animals. The information–theoretic analysis, outlined here, allows for a solid theoretical justification of this result. Moreover, the information–utility constraints can be used to determine the optimal dynamics by controlling parameter β (or the *temperature* parameter defined as $T \equiv \frac{1}{\beta}$).

It has been shown earlier how the entropy feedback from the posterior probability can be used to control β in the ACT–R architecture, which significantly improves cognitive models of human and animal learning [3]. A similar stochastic control has been used to implement optimal learning and adaptation of agents in stochastic environments [4]. The next section describes how such a stochastic process was implemented in our system of CAs of fLIF neurons, and how it is used to learn the connections between different CAs and modules. We shall also discuss biological plausibility of this meta–process.

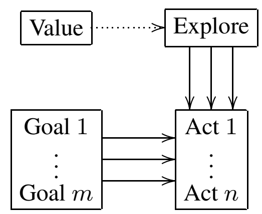
3.2 Stochastic control in cell assemblies

In the problems of learning agents, one often considers the set X of input patterns (e.g. describing the environment or the goals) and the set Y of actions of the agent. In our system, these sets can be represented by two modules, Goals and Actions, with CAs in the first module representing the input patterns (i.e. goals) and CAs in the second module representing different acts:



Our aim of learning the connections between these two modules can be described as learning some binary relation $R \subset X \times Y$. In fact, this is similar to defining a preference relation on set $\Omega = X \times Y$. Indeed, if some pairs (x, y) are preferred to the others, then given $x \in X$, there is a preference relation on Y. Moreover, if the agent has a preference relation on $\Omega = X \times Y$, then obviously it has to learn $R \subset X \times Y$ that corresponds to this preference relation.

Initially, there are excitatory connections from every CA in module X to all CAs in module Y, which means that all pairs (x, y) are equally preferred (i.e. indifference) and given goal $x \in X$, any action $y \in Y$ can be triggered. However, due to Hebbian learning, the connection $x \to y$ is reinforced if a particular pair of CAs ignite together, giving the pair a higher chance to ignite together in the future. Thus, simply by virtue of Hebbian learning, the system can learn eventually some random preference relation. The meta–process is designed to support the learning only of a particular preference relation, and it involves two additional modules: Explore and Value.



The purpose of the Explore module is to randomise the activity of the Action module. The Explore module contains cells that can be active without any external stimulation due to spontaneous activation. The connectivity and the parameters of the cells in the module are such that the activation can support itself. The cells in the Explore module send excitatory signals to all CAs in the Action net, and the weights of these connections do not change. Thus, the activity in the Explore module can trigger randomly any CA in the Action module, and this process does not have a memory. The activity of the Explore module implements the effect of the temperature parameter $T = \frac{1}{\beta}$ in equation (2).

The purpose of the Value module is to represent the values of the utility function — higher activity in the Value module corresponds to higher utility values $u = u(x, y)$. The input of the module can be configured according to the application. For example, it may receive inputs from the environment so that the activity of the Value module represents the agent's preference relation on the states of the environment. In the simplest case of a binary utility function (i.e. the utility has only two values corresponding to a success or failure), the Value module should have only two distinct states (on or off). For example, the module may ignite if the change of the environment is recognised as positive.

The Value module sends inhibitory connections to the Explore module, so that high activity of the Value cells may shut down the activity in the Explore module. As a result, any CA that has been ignited in the Action module will persist until it is shut down by another Action CA. The latter may ignite if the input from the Goal module changes or if the activity of the Explore module resumes. This connectivity implements a very simple yet effective learning scheme. If a particular goal–action pair (x, y) results in a high utility value, then high activity of the Value module inhibits the Explore module, and the responsible goal–action pair is allowed to persist longer. Since x and y co-fire longer than x and $\bar{y}$ (where $\bar{y}$ is a different CA than y) the $x \to y$ connection increases relative to the $x \to \bar{y}$ connection due to Hebbian learning.

Because the meta–process supports strengthening of the connections between the goal–action pairs corresponding to high utility values, the system learns the preferred binary relation $R \subset X \times Y$. As a consequence, the average activity of the Value module should increase with time, while the activity of the Explore module should decrease. This dynamic corresponds to an increase of the expected utility value $(p, u) = U$, and the decrease of the temperature parameter $T = \frac{1}{\beta}$ making the system less random and more deterministic.

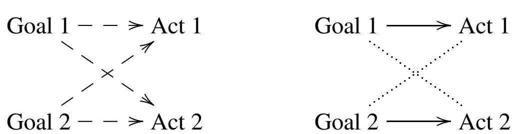
The process of learning the binary relation $R \subset X \times Y$ favouring high utility values results in a transition from a stochastic system to an almost deterministic rule–based system. The process of learning the connections $x \to y$ between the CAs can be seen as the emergence of 'if–then' rules, where the conditions are represented by CAs in one module and the actions by CAs in another.

4 EXPERIMENTAL EVALUATION

The working of the described meta–process has been implemented and tested in our system based on fLIF neurons, and here we report its performance in a fairly simple experiment. The code of the system and the described below experiment is available at
`http://www.cwa.mdx.ac.uk/CABot/CANT.html`

4.1 Learning dichotomies

In this simple experiment, there are two CAs in the Goal module (goal 1, goal 2) and two CAs in the Action module (act 1, act 2). Each module consisted of 800 cells, with 400 cells in each CA. The modules were set up with connections with low weights from every goal CA to all action CAs, shown by four dashed arrows on the left diagram below. The task was to learn two rules, shown by two solid arrows on the right diagram, by increasing the connection weights.



The training procedure consisted of a random presentation of an input pattern activating one of the goal CAs every 100 cycles. It takes on average 10–20 cycles for one of the action CAs to ignite. If the correct action is selected, then the activation of the Value module inhibits the Explore module after another 10–20 cycles, and the activities of the goal and action CAs persist until a new pattern is presented. Otherwise, if an incorrect action is selected, the activity from the Explore module causes another action CA to ignite after approximately another 10–20 cycles.

4.2 Results and analysis

Figure 1 shows the proportion of the correct actions selected (vertical axis) as a function of cycle number (horizontal axis). The chart shows the results of five similar experiments. One can see that the system initially makes only half of the choices correctly. After 3000 cycles, the proportion of correct choices increases to 70–90%. Note that the goal may change up to 10 times per 1000 cycles (every 100 cycles). Because the goal sequence was randomly generated in each experiment, there is a variance in the results represented by different curves on Figure 1. The increase of the probability of success corresponds to an increase in the expected utility value $(p, u) = U$.

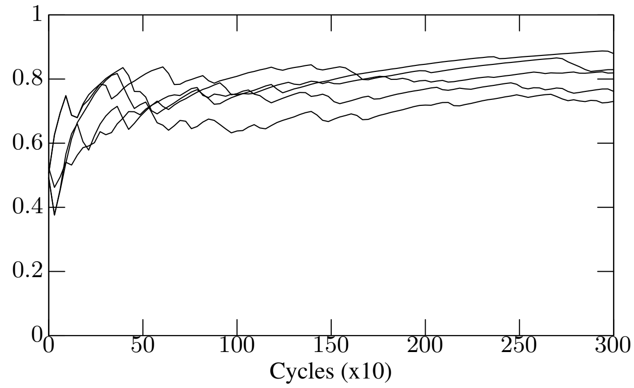


Figure 1. The proportion of correct action choices (ordinate) as a function of cycles (abscissa). The curves represent results of different trials.

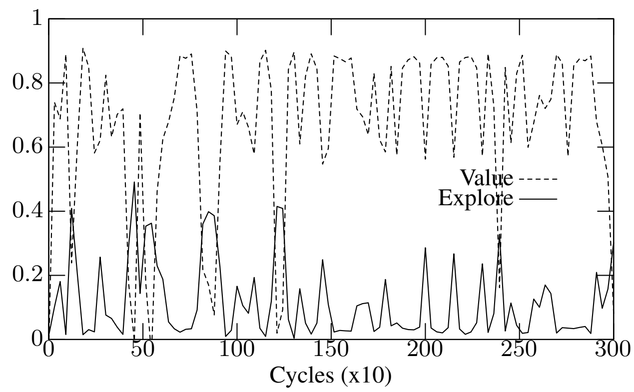


Figure 2. Activities of the Value and Explore modules in one experiment.

Figure 2 shows the percentage of neurons firing per cycle in the Value and the Explore modules in one of the experiments. One can clearly see that the activities anticorrelate. An increase in the Value module coincides with the decrease of the Explore module activity. More significantly, the chart shows that the average activity of the Value module increases as learning progresses, while the average activity of the Explore module decreases. As expected, this dynamic corresponds to the optimal dynamics of U and parameter $T = \frac{1}{\beta}$, where $\beta = \beta(U)$ is an increasing function defined by equation (4).

Because learning of the connections between the correct pairs of CAs depends on the differences between the times the 'correct' and 'incorrect' CAs persist in the system, the parameters controlling the dynamics of CAs in the modules may significantly influence the effect of the meta–process and the ability of the system to learn. For example, the values of the fatigue and fatigue recovery rates of the cells influence the persistence of the CAs as well as how rapidly one CA may extinguish another. Another important parameter is the connectivity of the cells in the module. The networks in the system are sparsely connected, and the average number of cells each cell is connected to can also significantly contribute to the behaviour of the CAs. The learning rate parameter of the Hebbian learning rule can also significantly influence the performance of the system. If the rate is too high, then binding of an incorrect pair of CAs may occur before the meta–process has its effect.

5 CONCLUSION

Computational learning theory has advanced greatly during recent decades, and there are excellent examples of connectionist and symbolic learning systems. Yet it is not clear how biological cognition combines these quite different approaches in one system. This question has been partially resolved by the ACT–R cognitive architecture [1], which uses a hybrid approach and combines the symbolic system with sub–symbolic computations based on statistical learning principles. In this work, we attempt to close the gap from the opposite direction. By using cell assemblies (CAs) as representations of symbols, we achieve a level of control in a complex system sufficient to implement symbolic algorithms. One of the problems that remains difficult to solve is how the connections between different and quite remote CAs can be learned in this system, the focus of the current paper.

The solution proposed is based on a stochastic meta–process that randomises the activation of the system according to the utility of its experience. This method has many similarities with the reinforcement learning algorithms, where randomisation is used to control exploration [12, 22], and with the adaptive networks where the reward signals were used to train artificial neurons [2, 21]. Here we have demonstrated how such a process can be implemented in a sparsely connected system of fLIF neurons, where CAs can be employed for symbolic–like processing. The implementation is inspired by earlier cognitive modelling work, where entropy feedback was used to control the stochastic learning in ACT–R, significantly improving models of action selection in human subjects and animals [3]. Information–theoretic analysis suggests that such a control corresponds to optimisation with information constraints.

Finally, recent studies in the neuroscience of exploratory behaviour suggest that the method proposed may have some biological plausibility. In particular, a recent study failed to identify conclusively any specific area of the brain correlated with the exploration function, and the model based on the Gibbs distribution was proposed as the most plausible [5]. Some researchers have speculated about the role of tonically active cholinergic neurons in the basal ganglia and striatal complex [6]. These neurons account for a small proportion of the connections, and they are quite uniform and non-topographic. It was suggested that these neurons may play the role of stochastic noise. Interestingly, their activation is reduced when the reward path is activated. This idea has remarkable parallels with the functioning of the Value and Explore modules in our system. Because learning occurs throughout the brain, it is possible that similar meta–processes exist in various areas of the central nervous system.

Our project is developing towards a complex system where many modules are combined together, implementing very different information processing functions. All the modules, however, are based on the same biologically inspired paradigm — cell assemblies of fLIF neurons. The implementation of the stochastic meta–learning process to allow rule acquisition in our system is an important step in its evolution, and the development of a biologically plausible mechanism creates new opportunities for the project as well as our understanding of biological cognition.

ACKNOWLEDGEMENTS

This work was supported by EPSRC grant EP/DO59720. We would like to thank the referees for their comments which helped improve this paper.

REFERENCES

[1] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, 'An integrated theory of the mind', *Psychological Review*, **111**(4), 1036–1060, (2004).

[2] A. G. Barto and P. Anandan, 'Pattern–recognizing stochastic learning automata', in *IEEE Transactions on Systems, Man and Cybernetics*, volume 15, pp. 360–375, (1985).

[3] Roman V. Belavkin, *On Emotion, Learning and Uncertainty: A Cognitive Modelling Approach*, PhD thesis, The University of Nottingham, Nottingham, UK, August 2003.

[4] Roman 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., Max Bramer, Frans Coenen, and Tony Allen, volume XXII of *Research and Development in Intelligent Systems*, pp. 305–316, Cambridge, (December 2005). BCS, Springer. ISBN 1-84628-225-X.

[5] N. D. Daw, J. P. O'Doherty, Peter Dayan, B. Seymour, and R. J. Dolan, 'Cortical substrates for exploratory decisions in humans', *Nature*, **441**(7095), 876–879, (2006).

[6] Richard Granger, 'Engines of the brain: The computational instruction set of human cognition', *AI Magazine*, **27**(2), 15–32, (July 2006).

[7] Donald O. Hebb, *The Organization of Behavior*, John Wiley & Sons, New York, 1949.

[8] Alan Lloyd Hodgkin and Andrew Huxley, 'A quantitative description of membrane current and its application to conduction and excitation in nerve', *Journal of Physiology*, **117**, 500–544, (1952).

[9] John Hopfield, 'Neural networks and physical systems with emergent collective computational abilities', *Proceedings of the National Academy of Sciences of the USA*, **79**, 2554–8, (1982).

[10] Christian Huyck, 'Hierarchical cell assemblies', *Connection Science*, (2006).

[11] Christian Huyck and Roman 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., Danilo Fum, Fabio Del Missier, and Andrea Stocco, Trieste, Italy, (April 2006). Edizioni Goliardiche.

[12] Leslie Pack Kaelbling, Michael L. Littman, and Andrew W. Moore, 'Reinforcement learning: A survey', *Journal of Artificial Intelligence Research*, **4**, 237–285, (1996).

[13] S. Kaplan, M. Sontag, and E. Chown, 'Tracing recurrent activity in cognitive elements (trace): A model of temporal dynamics in a cell assembly', *Connection Science*, **3**, 179–206, (1991).

[14] S. Kullback, *Information Theory and Statistics*, John Wiley and Sons, 1959.

[15] W. Maas and C. Bishop, *Pulsed Neural Networks*, MIT Press, 2001.

[16] Warren McCulloch and Walter Pitts, 'A logical calculus of ideas immanent in nervous activity', *Bulletin of Mathematical Biophysics*, **5**, 115–133, (1943).

[17] Allen Newell, *Unified theories of cognition*, Harvard University Press, Cambridge, Massachusetts, 1990.

[18] Yoshio Sakurai, 'The search for cell assemblies in the working brain', *Behavioural Brain Research*, **91**, 1–13, (1998).

[19] Claude E. Shannon, 'A mathematical theory of communication', *Bell System Technical Journal*, **27**, 379–423 and 623–656, (July and October 1948).

[20] Ruslan Leontevich Stratonovich, 'O tsennosti informatsii', *Izvestiya of USSR Academy of Sciences, Technical Cybernetics*, **5**, 3–12, (1965). In Russian.

[21] Richard S. Sutton and Andrew G. Barto, 'Toward a modern theory of adaptive networks: Expectation and prediction', *Psychological Review*, **88**(2), 135–170, (1981).

[22] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*, Adaptive Computation and Machine Learning, MIT Press, Cambridge, MA, 1998.

[23] J. von Neumann and O. Morgenstern, *Theory of games and economic behavior*, Princeton University Press, Princeton, NJ, first edn., 1944.

[24] Abraham Wald, *Statistical Decision Functions*, John Wiley & Sons, New York, 1950.