Initial Results from the use of Learning Classifier Systems to Control *In Vitro* Neuronal Networks

Larry Bull
School of Computer Science
University of the West of England
Bristol BS16 1QY, U.K.
+44 (0)117 3283161
Larry.Bull@uwe.ac.uk

Ivan S. Uroukov
Faculty of Applied Sciences
University of the West of England
Bristol BS16 1QY, U.K.
+44 (0) 117 3283880
Ivan.Uroukov@uwe.ac.uk

ABSTRACT
In this paper we describe the use of a learning classifier system to control the electrical stimulation of cultured neuronal networks. The aim is to manipulate the environment of the cells such that they display elementary learning, i.e., so that they respond to a given input signal in a pre-specified way. Results indicate that this is possible and that the learned stimulation protocols identify seemingly fundamental properties of *in vitro* neuronal networks. Use of another learning scheme and simpler stimulation confirms these properties.

Categories and Subject Descriptors
I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods and Search – *backtracking, control theory, dynamic programming, graph and tree search strategies, heuristic methods, plan execution formation and execution, scheduling.*

General Terms
Algorithms, Measurement, Experimentation.

Keywords
Multi-Electrode Array, Unconventional Computation, XCS.

1. INTRODUCTION
There is growing interest in research into the development of hybrid wetware-silicon devices focused on exploiting their potential for non-linear media computing, particularly cultured neurons. The aim is to harness the as yet only partially understood intricate dynamics of *in vitro* neuronal networks to perform complex 'computations' (potentially) more effectively than with traditional AI architectures and to further the understanding of how nervous systems function. The area provides the prospect of radically new forms of machines and is enabled by improving Capabilities in cell culturing, neurobiology and wetware-silicon interfacing. Such systems also have many possible medical uses such as in prosthetics, the study of degenerative diseases, etc. The study of *in vitro* networks has the potential to discover the underlying behaviours of neurons since they are typically created from dissociated cells; the self-organizing characteristics of such cells become identifiable. Such networks have already been reported as being capable of simple learning, memory and other computation-like behaviours.

It is well-established that *in vitro* neuronal networks display a strong disposition to form synapse and sensitivity to electro-chemical stimulation. Shahaf and Marom [17] have highlighted these latter characteristics in their work with cultured rat neurons in commercially available multi-electrode hardware (Multichannel Systems Ltd. MEA-60, as shown in Figure 1). Some electrodes are designated as input sources and those remaining are monitored for recurring patterns in action potentials; such technology enables network/ensemble level analyses (e.g., [14]). They were able to demonstrate a simple form of supervised stimulus-response learning in the cultured networks such that a required response for a given input was obtained from a pre-determined neuron/electrode through timed stimulus removal. That is, drawing on ideas proposed during the 1940's by behavioural psychologists, they showed that with incremental single-step training, desired discrete output computations could be achieved from essentially randomly connected neuronal networks.

Shahaf and Marom's [ibid.] work is related to that by DeMarse et al. [6] who have used the same hardware to randomly control a simulated mobile robot, again with feedback from output to the inputs. That is, they have presented an approach to *in vitro* AI wherein the neuronal network exists within a feedback loop to its environment: the sensors of the simulated mobile robot are fed directly into the network and its responses fed to the robot's actuators. They report the emergence of a number of repeated spiking patterns during the control scenario.

Ruaro et al. [15] describe the use of neuronal networks for an image processing task. Here two spatial patterns are exposed to the network through appropriate electrode stimulation. They show that the response of the network to one pattern can be trained to be significantly higher than for the other.

It has also recently been shown that controlled pair-wise stimulation can be used to alter network response thereby indicating a rudimentary memory mechanism for *in vitro*

GECCO'07, July 7-11, 2007, London, England, United Kingdom.

networks; response to a given stimulus on one electrode alters if another has been stimulated within a time window (e.g., [21]).

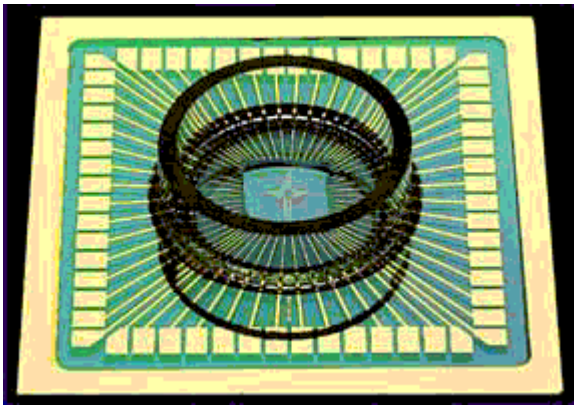


Figure 1. Multi-electrode array technology, showing the dish in which the neuronal network grows. Image from Multi Channel Systems.

We are currently exploring the use of evolutionary computing techniques, particularly versions of Holland's Learning Classifier System (LCS) [9], to manipulate the stimulation of *in vitro* neuronal networks with the aim of shaping their behaviour. In this paper we present initial results from our approach.

2. A LEARNING CLASSIFIER SYSTEM

XCS [26] is a significant development of Holland's Learning Classifier System formalism and has been shown able to tackle many complex tasks effectively (see [2] for examples). It consists of a limited size (N) population [P] of classifiers (rules). Each classifier is in the form of "IF condition THEN action" (*condition*→*action*) and has a number of associated parameters. Conditions traditionally consist of a trinary representation, {0,1,#}, where the wildcard symbol facilitates generalization, and actions are binary strings.

On each time step a match set [M] is created. A system prediction is then formed for each action in [M] according to a fitness-weighted average of the predictions of rules in each action set [A]. The system action is then traditionally selected either deterministically or stochastically based on the fitness-weighted predictions (usually 0.5 probability per trial). If [M] is empty a covering heuristic is used which creates a random condition to match the given input and then assigns it to a rule for each possible action.

Fitness reinforcement in XCS consists of updating three parameters, ε, p and F for each appropriate rule; the fitness is updated according to the relative accuracy of the rule within the set in five steps:

i) Each rule's error is updated: $\varepsilon_j = \varepsilon_j + \beta(| P - p_j | - \varepsilon_j)$ where $0 \leq \beta \leq 1$ is a learning rate constant.

ii) Rule predictions are then updated: $p_j = p_j + \beta(P - p_j)$

iii) Each rule's accuracy κ_j is determined:

$$\kappa_j = \alpha(\varepsilon_0/\varepsilon)^{\nu} \text{ or } \kappa=1 \text{ where } \varepsilon < \varepsilon_0$$

ν, α and ε_0 are constants controlling the shape of the accuracy function.

iv) A relative accuracy κ_j' is determined for each rule by dividing its accuracy by the total of the accuracies in the action set.

v) The relative accuracy is then used to adjust the classifier's fitness F_j using the moyenne adaptive modifee (MAM) procedure: If the fitness has been adjusted $1/\beta$ times, $F_j = F_j + \beta(\kappa_j' - F_j)$. Otherwise F_j is set to the average of the values of κ' seen so far.

In short, in XCS fitness is an inverse function of the error in reward prediction, with errors below ε_0 not reducing fitness. The maximum $P(a_i)$ of the system's prediction array is discounted by a factor γ and used to update rules from the previous time step and an external reward may be received from the environment. Thus XCS exploits a form of Q-learning [25] in its reinforcement procedure.

A Genetic Algorithm (GA) [8] acts in action sets [A], i.e., niches. Two rules are selected based on fitness from within the chosen [A]. Two-point crossover is applied at rate χ and point mutations at rate μ. Rule replacement is global and based on the estimated size of each action set a rule participates in with the aim of balancing resources across niches. The GA is triggered within a given action set based on the average time since the members of the niche last participated in a GA (after [1]).

The intention in XCS is to form a complete and accurate mapping of the problem space through efficient generalizations. In reinforcement learning terms, XCS learns a value function over the complete state/action space. In this way, XCS represents a means of using temporal difference learning on complex problems where the number of possible state-action combinations is very large (other approaches have been suggested, such a neural networks – see [20] for an overview). The reader is referred to [4] for an algorithmic description of XCS and [3] for an overview of current formal understanding of its operations.

3. NEURONAL NETWORKS

The majority of *in vitro* studies of the electrophysiological properties of neuronal networks exploit either tissue slices or monolayer cell cultures. For example, all the research described in the introduction used monolayers, i.e., cells in a network grown across the surface of the multi-electrode array dish. However it has long been known that aggregated, i.e., 3-D, neuronal cell cultures exhibit properties that are remarkably similar to their *in vivo* counterparts. For example, early studies showed structures identical to hippocampal architecture [5][18] and Seeds [16] showed how the temporal biochemical differentiation of brain cell aggregates was very similar to that seen during development in mice, much more so than equivalent monolayer cultures. Indeed, the amount and type of cell differentiation was suggested to be the main difference between monolayer and aggregate cultures (e.g., [13] [22]).

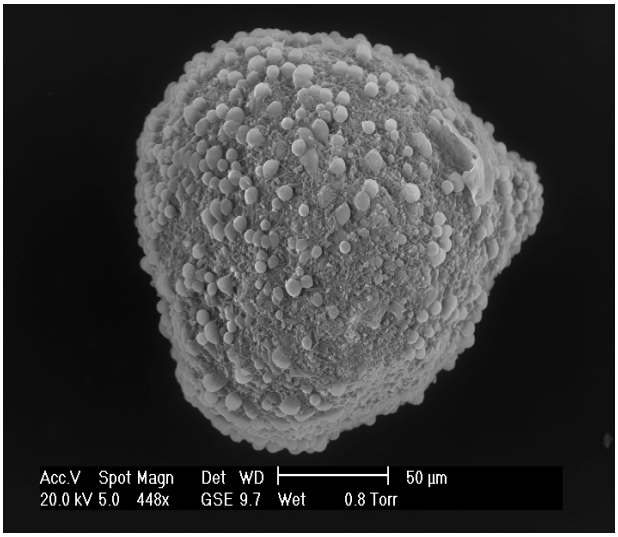


Figure 2. Scanning Electron Microscope image of DIV21 hen embryo aggregate neuronal culture.

Advances in cell culturing mean that it is now possible to differentiate neuronal and neuroglial cells obtained from ovoid primary cultures and maintain them for relatively long periods of time, typically several months. These organotypic cultures are derived from hen embryos at day 7 *in ovo*. We have recently described how the maturation of spontaneous spiking behaviour in aggregated cultures of such cells is typically very similar to that reported in monolayers of mammalian cortical cells [23]. However, response to simple stimulation has been shown able to cause an increase in the relative spiking frequency during maturation, typically up to around two times larger after fourteen days in culture (DIV14). This result indicates strong self-organizing processes within the neuronal networks of such aggregate cultures wherein networks containing mutual inhibition form under steady-state (unstimulated) conditions in such a way that external stimulation causes significant excitation within the structure. It is this feature we aim to explore further using evolutionary computation. Figure 2 shows an example of the aggregates used in this study. The reader is referred to [23] for details of the cell culturing protocol used here.

Multi-electrode arrays (MCS-2100, 3-D Multi Electrode Array, Multi Channel Systems MCS GmbH, Aspenhaustrasse 21, 72770 Reutlingen, Germany) with (3D-40 x 40 x 70 µm, spaced on 200 µm) pyramid shaped electrodes were used to record electrical activity of the aggregates. The multi-electrode (MEA) dish surface was modified with 10 µg/ml aqueous solution of Polymer Ethylene Imine (PEI) (Fluka Chemie AG, Buchs, Switzerland) under sterile conditions. The molecular weight of PEI varied between 0.610 and 1.010 according to product specifications. After the modification, two washing steps with demineralized (DEMI) water were undertaken before the plating of the aggregates.

The electrical recordings from the cell aggregates was performed with a 60 channels data acquisition system, where the sampling frequency of each channel was set to 25 kHz and the single channel amplification kept at 1200 with a digital resolution of 12 bits. At these conditions, data sampling of the input band of spikes within 5 kHz including a high pass 300Hz filter was performed in a way that was similar to other studies (e.g., [12][7]). The spikes were detected by threshold depending upon the standard deviation and the offset of noise. A set of data was monitored and raw signal, filtered signal and spikes chosen in order to perform fast and reliable recordings and analysis with the MC Rack software (Multi Channel Systems MCS, GmbH). The recorded data was written in the custom '*.mcd' MC Rack format and stored for further data analysis. The recorded signals were analyzed and the spike parameters extracted using the MC Rack software and analyzed with bespoke software as a post processing step.

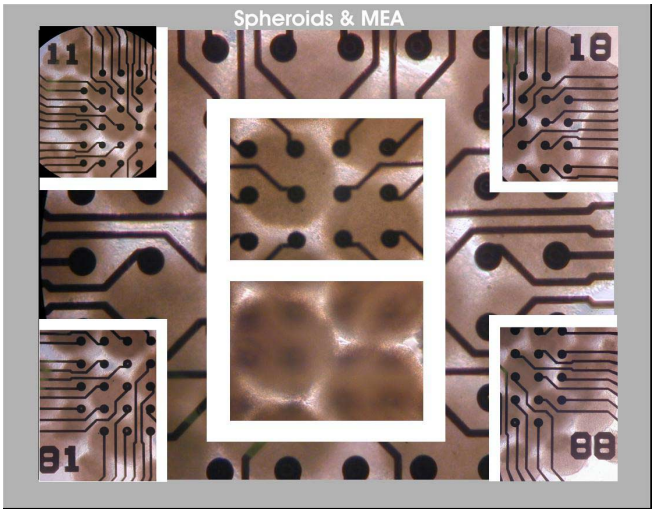


Figure 3. Phase contrast microscopy images taken at various optical magnifications and focal planes of aggregate cell cultures on multi-electrode array dish.

The stimulation of the aggregate cell culture neuronal networks was realized with the 8 channel programmed generator STG2008 (Multi Channel Systems MCS, GmbH). The stimulation protocol was created within the MC Stimulus II software (Multi Channel Systems MCS, GmbH). The elaborated protocol shared features of relevant published studies (e.g., [24]), consisting of a single sharp biphasic impulse of 300 µs and voltages between 300 mV and 2000 mV each phase per sweep of 1 sec.

4. XCS CONTROL

In the current study XCS has been applied to the control of the electrical stimulation of the neuronal networks in the following

way. Firstly, the average spontaneous spiking frequency of a chosen aggregate network is ascertained over a 300 second window. Typically, an individual aggregate covers three or four electrodes in a dish as shown in Figure 3, one or two of which will show a suitably good connection into the neuronal network therein, i.e., spikes will be detected of the kind shown in Figure 4. The standard deviation in the spiking frequency is also calculated over the window. The task of the XCS controller is then to cause the chosen neuronal network to reply to the simple stimulus described above with a spiking frequency of the spontaneous mean plus two standard deviations; a significant increase in typical spiking frequency is required under stimulation.

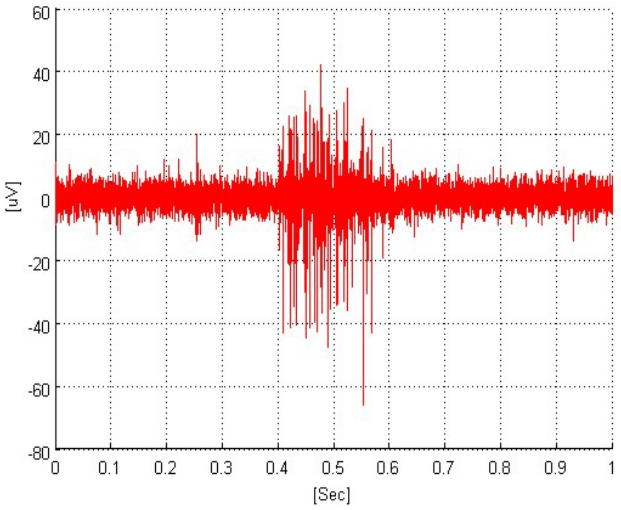


Figure 4. Example spiking behaviour recorded on a single electrode.

The input to the XCS on each cycle is the spiking frequency of the neuronal network averaged over the last three seconds and the length of time the stimulus was applied. The first number is presented as a fraction of the maximum spiking frequency observed under the 300 seconds of spontaneous behaviour and the second as a fraction of the maximum allowed stimulation time of 600 seconds. The XCS returns one of three actions: to double, halve or maintain the current stimulation time. A reward of 500 is given if the spiking frequency increased over the last stimulation period compared to that immediately prior and a reward of 1000 is given if the target spiking frequency, or one greater, was achieved.

Following [17] we allow a 300 second rest period between applications of the stimulus and truncate the maximum duration of stimulation to be 600 seconds. Thus 300 seconds after the last stimulation period, the XCS controller is given the last recorded spiking frequency of the neuronal network under stimulation, as a three-point running average, and the amount of time for which the stimulus was applied that caused the response. It then adjusts or maintains the stimulus duration for the coming cycle. For the initial cycle a stimulation period of 60 seconds is used.

Hence the XCS is presented with an environmental input consisting of two real numbers scaled between 0.0 and 1.0; the condition part of the classifiers is encoded as un-ordered pairs of real numbers in the range [0, 1], one pair for each environmental input (after [19]). A pair is considered to match the corresponding input value if one of the pair is smaller or equal to the target, and the other is larger or equal. The action of the classifier is an integer.

The mutation operator is altered from that in XCS as described above to deal with the new representation. Mutation in the case of the real numbers of the condition is effected either by the addition or subtraction of either a small number drawn from a Gaussian distribution centred on the current value, or a fixed small change (here 0.1). Action mutation is by picking an integer from the set {0,1,2} at random, such that the chosen action is different from the current one.

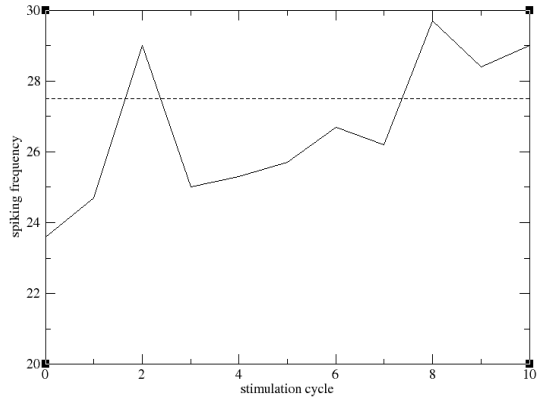
In the initial population, classifier conditions are created randomly in the range [0,1]. During cover, the current environmental input e is used as a centre and two values are created in the range $[e\text{-}C_{max}, e\text{+}C_{max}]$, where C_{max} is 0.1.

One further change is made to the standard XCS described above: roulette wheel action selection is used on explore trials rather than random selection since this is more appropriate to the on-line learning scenario.

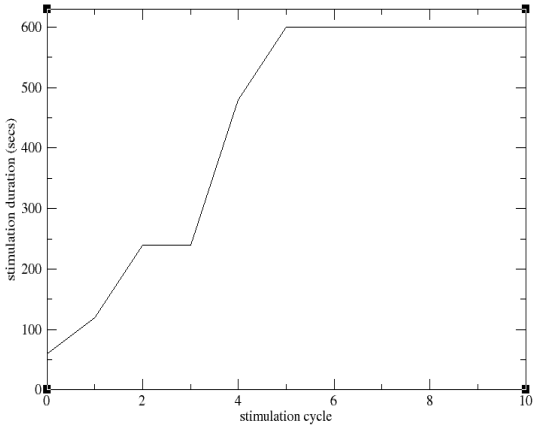
The XCS parameters used were typical for those in the literature: N=3000, β=0.2, μ=0.04, χ=0.8, θ_{del}=20, δ=0.1, ε_0=10, α=0.1, ν=5.0, θ_{mna}=3, θ_{GA}=2, $\rho_I = \varepsilon_I = F_I$=10.0. The reader is referred to [4] for a full description of these parameters. The 20 cell cultures used were all in the range of 20-30 days old *in vitro*.

5. RESULTS

After a number of experiments it became clear that the XCS was only able to alter a neuronal network's behaviour in roughly a third of cases. Figure 5 shows an example where it was able to cause the required spiking response to the stimulus. As can be seen, and as was typical here, the XCS controller achieves this by increasing the duration for which the stimulation is applied. However Figure 6 shows a case where no significant change in spiking appears to have occurred regardless of how the XCS adjusts the network's stimulation. In other cases the average spiking response decreases during the experiment regardless of stimulation duration (not shown). These figures show both the exploration and exploitation trial behaviour of XCS, i.e., the actual on-line stimulation as experienced by the neuronal networks. This is in contrast to the more typical presentation of exploit trials only (e.g., after [26]).



(a)

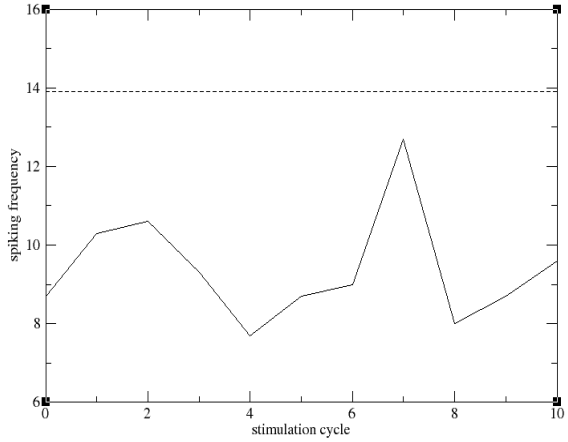


(b)

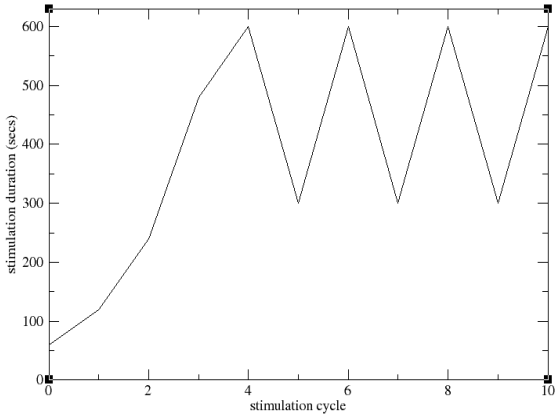
Figure 5. Example learning behaviour under XCS control, showing the spiking frequency response becoming repeatedly higher than the target indicated by the dashed line (a) and how XCS altered the stimulus application time to achieve this (b).

Given these findings we implemented the aforementioned behaviour shaping/learning protocol of Shahaf and Marom [17] for comparison. This scheme, in contrast to the more widespread consideration of neuromodulatory reward mechanisms for learning, is inspired by the work of behaviourists such as Clark Hull (e.g., [10]) in the 1940's. Known as the Stimulus Regulation Principle (SRP) it proposes that reward and hence learned behaviour is achieved through the removal of the driving stimulus. That is, neurons cease a constant alteration to their connectivity when the driving stimulus is removed and hence the behaviour becomes fixed; no other mechanism, i.e., no neuromodulator, is required for such (low-level) learning.

In our implementation the target spiking frequency was again the mean plus two standard deviations recorded under spontaneous behaviour for 300 seconds. The same stimulus was applied as before and removed either when the required spiking response was obtained (as a running average over the last 3 seconds, as before) or if 600 seconds had elapsed. Again, a 300 second rest period between applications was allowed.



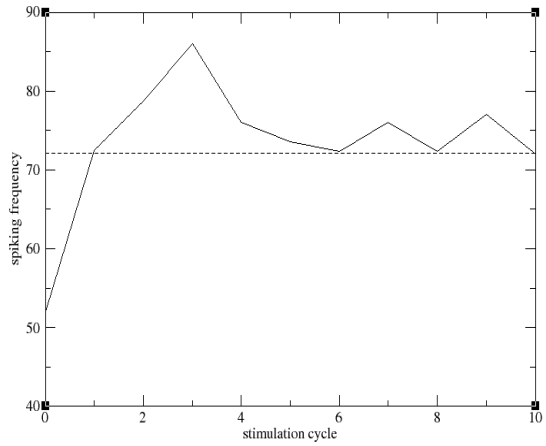
(a)



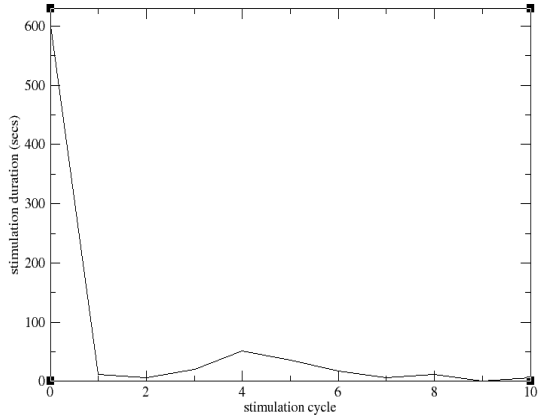
(b)

Figure 6. Example unsuccessful learning behaviour under XCS control, showing the spiking frequency response never rising to the target indicated by the dashed line (a) and how XCS altered the stimulus application time (b).

Figure 7 shows a successful experiment akin to those reported by Shahaf and Marom. Here the amount of time the stimulus must be applied before the required spiking frequency is seen rapidly decreases until it is consistently obtained almost immediately with every application. To our knowledge this represents the first reproduction of the work by Shahaf and Marom. However we again found that such results occurred only about a third of the time. Figure 8 shows an example where the target frequency is never seen, i.e., the stimulation always remains applied for the maximum of 600 seconds, and the spiking frequency drops over time. Examples with no significant change were again also seen (not shown).



(a)

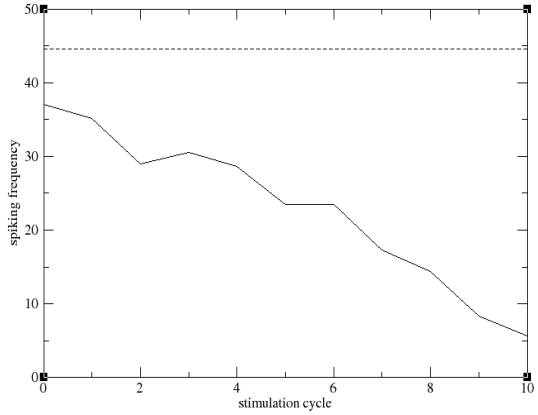


(b)

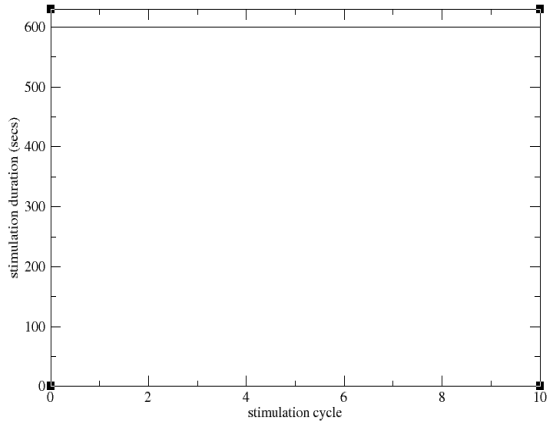
Figure 7. Example learning behaviour under SRP control, showing the spiking frequency response becoming repeatedly higher than the target indicated by the dashed line (a) and how the stimulus was applied/removed to achieve this (b). The spiking frequency shown is the last recorded on a given cycle.

6. CONCLUSIONS

The results from using XCS, and then an SRP-inspired protocol, to induce learning indicate three possible rudimentary responses to simple stimulation from such *in vitro* 3-D neuronal networks: *excitation,* giving the potential for significant increases in typical spiking behaviour; *inhibition*, wherein spiking behaviour decreases due to stimulation; and *unchanging*, i.e., no significant shift in spiking behaviour over spontaneous behaviour is seen due to the stimulus. We found that each such behaviour was equally likely and that neither learning protocol could affect the underlying behaviour of a given neuronal network.

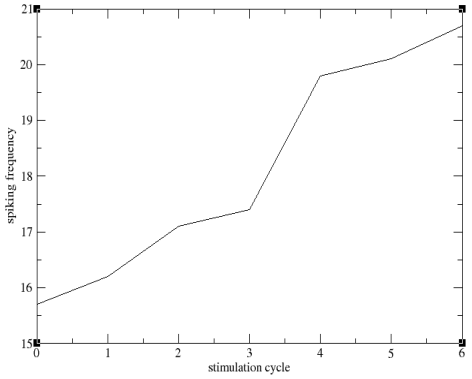


(a)

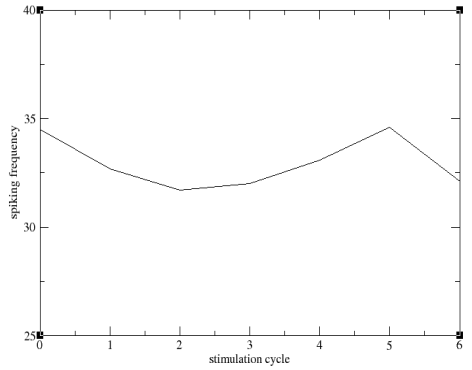


(b)

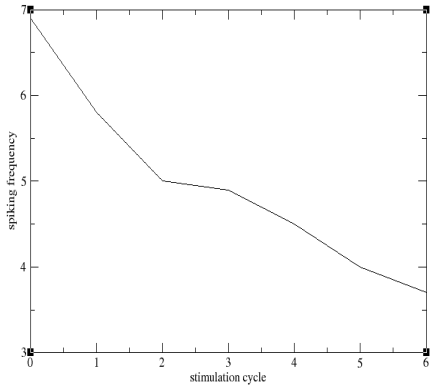
Figure 8. Example unsuccessful learning behaviour under SRP control, showing the spiking frequency response falling away from the target indicated by the dashed line (a) and how the stimulus was therefore constantly applied throughout (b). The spiking frequency shown is the last recorded on a given cycle.



(a)



(b)



(c)

Figure 9. Showing examples of the three responses obtained under the simple stimulation protocol wherein the stimulus is constantly applied for 600 seconds before a rest of 300 seconds.

To examine how strong these underlying behaviours might be we also stimulated cultures in a very simple way, that is, constantly for periods of 600 seconds, followed by a 300 second rest period. Figure 9 shows how the same three responses were again obtained.

Jimbo et al. [11], using monolayer cultures of mammalian cortex, reported simple stimulation to an electrode could induce either an excitatory or inhibitory response at other electrodes - "pathway-dependent plasticity". Our results therefore suggest the same is also true for aggregate neuronal cell cultures but that a third class of behaviour is possible. It is not clear at this time whether such behaviour is typical in monolayers.

We are currently exploring the use of XCS to elicit more subtle responses to stimulation from such neuronal networks and to use them for computation.

7. ACKNOWLEDGMENTS

This work was support under EPSRC Grant No. GR/T11029.

8. REFERENCES

[1] Booker, L.B. (1989) Triggered Rule Discovery in Classifier Systems. In J.D. Schaffer (ed) *Proceedings of the Third International Conference on Genetic Algorithms.* Morgan Kaufmann, pp265-274.

[2] Bull, L. (2004)(ed) *Applications of Learning Classifier Systems*. Springer.

[3] Bull, L. & Kovacs, T. (2005)(eds) *Foundations of Learning Classifier Systems*. Springer.

[4] Butz, M. & Wilson, S.W. (2002) An Algorithmic Description of XCS. *Soft Computing* 6(3): 144-153.

[5] DeLong, G.R. (1970) Histogenesis of fetal mouse isocortex and hippo-campus in reaggregating cell cultures. *Dev. Biol.* 22:563-583.

[6] DeMarse, T.B., Wagenaar, D.A., Blau, A.W. & Potter, S.M. (2001) The Neurally Controlled Animat: Biological Brains acting with Simulated Bodies. *Autonomous Robotics* 11: 305-310.

[7] Guillory, K.S. & Norman, R.A. (1999) A 100-channel System for Real Time Detection and Storage of Extracellular Spike Waveforms. *Journal of Neuroscience Methods* 91: 21-29.

[8] Holland, J.H. (1975) *Adaptation in Natural and Artificial Systems*. University of Michigan Press.

[9] Holland, J.H. (1986) Escaping Brittleness. In R.S. Michalski, J.G. Carbonell & T.M. Mitchell (eds) *Machine Learning: An Artificial Intelligence Approach, 2*. Morgan Kauffman, pp48-78.

[10] Hull, C. (1943) *Principles of Behaviour*. Appleton-Century-Crofts.

[11] Jimbo, Y., Tateno, T. & Robinson, H. (1999) Simultaneous Induction of Pathway-Specific Potentiation and Depression in Networks of Cortical Neurons. *Biophysics Journal* 76(2): 670-678.

[12] Jimbo, Y., Kawana, A., Parodi, P. & Torre, V. (2000) The Dynamics of a Neuronal Culture of Dissociated Cortical Neurons of Neonatal Rats. *Biol. Cybern.* 83: 1-20.

[13] Moscona A., (1961) Rotation mediated histogenic aggregation of dissociated cells. *Exp. Cell Res.* 22: 455–475.

[14] Potter, S.M. (2001) Distributed Processing in Cultured Neuronal Networks. In M. Nicolelis (ed) *Progress in Brain Research* vol. 130, pp1-14.

[15] Ruaro, M.E., Bonifazi, P. & Torre, V. (2005) Toward the Neurocomputer: Image Processing and Pattern Recognition with Neuronal Cultures. *IEEE Trans. on Biomedical Engineering* 52(3): 371-383.

[16] Seeds, N.W. (1971) Biochemical Differentiation in Reaggregating Brain Cell Culture. *Proc. Nat. Acad. Sci. USA* 68(8): 1858-1861.

[17] Shahaf, G. & Marom, S. (2001) Learning in networks of cortical neurons. *Journal of Neuroscience* 21 (22):8782-8788.

[18] Sidman, R.L. (1970). In F.O. Schmitt (Ed.) *The Neurosciences Second Study Program.* Rockfeller University Press, pp100.

[19] Stone, C & Bull, L. (2003) For Real! XCS with Continuous-Valued Inputs. *Evolutionary Computation* 11(3): 299-336

[20] Sutton, R. & Barto, A. (1998) *Reinforcement Learning*. MIT Press.

[21] Takayam, Y. & Jimbo, Y. (2006) Modification of Evoked Responses Induced by Correlated Stimuli in Cultured Cortical Networks. In *Proceedings of the 5th International Meeting on Substrate-Integrated Micro Electrode Arrays*. BIOPRO, pp22-25.

[22] Trapp B. D., Honneger, P., Richelson, E. & Webster, H. deF. (1979) Morphological Differentiation of Mechanically Dissociated Fetal Rat Brain in Aggregating Cell Cultures. *Brain Research* 160:117-180.

[23] Uroukov, I., Ma, M., Bull, L. & Purcell, W. (2006) Electrophysiological Measurements in 3-Dimensional *In Vivo*-Mimetic Organotypic Cell Cultures: Preliminary Studies with Hen Embryo Brain Spheroids. *Neuroscience Letters* 404: 33-38.

[24] Wagenaar, D., Madhavan, R., Pine, J. & Potter, S.M. (2005) Controlling Bursting in Cortical Cultures with Closed-Loop Multi-Electrode Stimulation. *Journal of Neuroscience* 25(3): 680–688.

[25] Watkins, C.J. (1989) Learning from Delayed Rewards. Ph.D. Thesis, Cambridge University.

[26] Wilson, S.W. (1995) Classifier Fitness Based on Accuracy. *Evolutionary Computation* 3(2):149-17.