Towards Unconventional Computing through Simulated Evolution: Control of Nonlinear Media by a Learning Classifier System

Larry Bull*,**
University of the West of England

Adam Budd**
University of the West of England

Christopher Stone**
University of the West of England

Ivan Uroukov†
University of the West of England

Ben de Lacy Costello†
University of the West of England

Andrew Adamatzky**
University of the West of England

Abstract We propose that the behavior of nonlinear media can be controlled automatically through evolutionary learning. By extension, forms of unconventional computing (viz., massively parallel nonlinear computers) can be realized by such an approach. In this initial study a light-sensitive subexcitable Belousov-Zhabotinsky reaction in which a checkerboard image, composed of cells of varying light intensity projected onto the surface of a thin silica gel impregnated with a catalyst and indicator, is controlled using a learning classifier system. Pulses of wave fragments are injected into the checkerboard grid, resulting in rich spatiotemporal behavior, and a learning classifier system is shown to be able to direct the fragments to an arbitrary position through dynamic control of the light intensity within each cell in both simulated and real chemical systems. Similarly, a learning classifier system is shown to be able to control the electrical stimulation of cultured neuronal networks so that they display elementary learning. Results indicate that the learned stimulation protocols identify seemingly fundamental properties of in vitro neuronal networks. Use of another learning scheme presented in the literature confirms that such fundamental behavioral characteristics of a given network must be considered in training experiments.

Keywords
Chemical computing, genetic algorithm, neuronal computing, reinforcement learning

1 Introduction

There is growing interest in research into the development of hybrid wetware-silicon devices focused on exploiting their potential for nonlinear computing. The aim is to harness the as yet only partially understood intricate dynamics of nonlinear media to perform complex computations, (potentially) more effectively than with traditional architectures, and to further the understanding of how such systems function. The area provides the prospect of radically new forms of machines and is enabled by improving capabilities in wetware-silicon interfacing. We are developing an approach by which networks of nonlinear media—reaction-diffusion systems and in vitro neuronal networks—can be

* Corresponding author.
** Faculty of Computing, Engineering & Mathematics, University of the West of England, Coldharbour Lane, Frenchay, Bristol BS16 1QY, U.K. E-mail: larry.bull@uwe.ac.uk
† Faculty of Applied Sciences, University of the West of England, Coldharbour Lane, Frenchay, Bristol BS16 1QY, U.K.

produced to achieve user-defined computation in a way that allows control of the media used. Evolutionary algorithms (e.g., [33]) are used to design the appropriate networks by searching a defined behavioral space to create a computing resource capable of satisfying one or more given objectives. In this article we begin by examining a Belousov-Zhabotinsky (BZ) [54] reaction-diffusion system in which the networks are created with light, and we present initial results from the general control-programming methodology. We then apply the same approach to the control of a 3D form of cultured neuronal network.

Excitable and oscillating chemical systems have been used to perform a number of computational tasks [1] such as implementing logical circuits [42, 48], image processing [30], shortest-path problems [41], and memory [35]. In addition, chemical diodes [2], coincidence detectors [16], and transformers where a periodic input signal of waves may be modulated by the barrier into a complex output signal depending on the gap width and frequency of the input [40] have all been demonstrated experimentally.

A number of experimental and theoretical constructs utilizing networks of chemical reactions to implement computation have been described. These chemical systems act as simple models for networks of coupled oscillators such as neurons, circadian pacemakers, and other biological systems [28]. Over 30 years ago the construction of logic gates in a bistable chemical system was described by Rossler [36]. Ross and coworkers [18, 19] produced a theoretical construct suggesting the use of "chemical" reactor systems coupled by mass flow for implementing logic gates, neural networks, and finite-state machines. In further work Hjelmfelt et al. [17, 20] simulated a pattern recognition device constructed from large networks of mass-coupled chemical reactors containing a bistable iodate–arsenous acid reaction. They encoded arbitrary patterns of low and high iodide concentrations in the network of 36 coupled reactors. If the network is initialized with a pattern similar to the encoded one, then errors in the initial pattern are corrected, bringing about the regeneration of the stored pattern. However, if the pattern is not similar, then the network evolves to a homogeneous state signaling nonrecognition.

In related experimental work Laplante et al. [31] used a network of eight bistable mass-coupled chemical reactors (via 16 tubes) to implement pattern recognition operations. They demonstrated experimentally that stored patterns of high and low iodine concentrations could be recalled (stable output state) if similar patterns were used as input data to the programmed network. This highlights how a programmable parallel processor could be constructed from coupled chemical reactors. This chemical system has many properties similar to parallel neural networks. In other work, Lebender and Schneider [32] described methods of constructing logical gates using a series of flow-rate-coupled continuous-flow stirred tank reactors (CSTRs) containing a bistable nonlinear chemical reaction. The minimal bromate reaction involves the oxidation of cerium(III) (Ce^{3+}) ions by bromate in the presence of bromide and sulfuric acid. In the reaction the Ce^{4+} concentration state is considered as "0" or "false" ("1" or "true") if a given steady state is within 10% of the minimal (maximal) value. The reactors were flow-rate coupled according to rules given by a feedforward neural network run using a PC. The experiment is started by feeding in two "true" states to the input reactors and then switching the flow rates to generate "true"-"false", "false"-"true", and "false"-"false". In this three-coupled-reactor system, AND (output "true" if inputs are both high Ce^{4+}, "true"), OR (output "true" if one of the inputs is "true"), NAND (output "true" if one of the inputs is "false"), and NOR (output "true" if both of the inputs are "false") gates could be realized. However, to construct XOR and XNOR gates two additional reactors (a hidden layer) were required. These composite gates are produced by interlinking AND and OR gates and their negations. In this work, coupling was implemented by computer, but the authors suggested that true chemical computing of some Boolean functions may be achieved by using the outflows of reactors as the inflows to other reactors, that is, serial mass coupling.

As yet no large-scale experimental network implementations have been undertaken, mainly due to the complexity of analyzing and controlling many reactors. That said, there have been many experimental studies carried out involving coupled oscillating and bistable systems [3, 4, 9, 10, 23, 44]. The reactions are coupled together either physically by diffusion or an electrical connection, or

chemically by having two oscillators that share a common chemical species. The effects observed include multistability, synchronization, in-phase and out-of-phase entrainment, amplitude or oscillator death, the cessation of oscillation in two coupled oscillating systems, and the reverse, *rhythmogenesis*, in which coupling two systems at steady state causes them to start oscillating [13].

In this article we adapt a system described by Wang et al. [51] and explore the computational potential based on the movement and control of wave fragments. In the system they describe, Gaussian noise (where the mean light level is fixed at the subexcitable threshold of the reaction) in the form of light is projected onto a thin layer of the light-sensitive analogue of the BZ reaction. This was observed to induce wave formation and subsequently *avalanche behavior*, whereby a proliferation of open-ended excitation wave fragments occurred. Interestingly, calcium waves induced in networks of cultured glial cells [27] display similar features to the ones identified in this chemical system, which the authors postulated may provide a possible mechanism for long-range signaling and memory in neuronal tissues.

The study of in vitro neuronal networks has the potential to discover the underlying behaviors of neurons, since such networks 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 behaviors.

It is well established that in vitro neuronal networks display a strong disposition to form synapses and sensitivity to electrochemical stimulation. Shahaf and Marom [39] have highlighted these characteristics in their work with cultured rat neurons in commercially available multi-electrode hardware (Multichannel Systems Ltd. MEA-60). Some electrodes are designated as input sources, and the rest are monitored for recurring patterns in action potentials; such technology enables large-scale network- or ensemble-level analyses. Shahaf and Marom 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 predetermined neuron (electrode) through timed stimulus removal. Drawing on ideas proposed during the 1940s by behavioral 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 [39] work is related to that of DeMarse et al. [12], who used the same hardware to randomly control a simulated mobile robot, again with feedback from the output to the inputs. They 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 process.

Ruaro et al. [37] 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 the response to the other. It has also recently been shown that controlled pairwise stimulation can be used to alter network response, thereby indicating a rudimentary memory mechanism for in vitro networks; the response to a given stimulus on one electrode alters if another has been stimulated within a time window (e.g., [46]).

Machine learning techniques, such as evolutionary algorithms (EAs) (e.g., [37]) and reinforcement learning (RL) (e.g., [45]), are being increasingly used in the design of complex systems. Example applications include data mining, time series analysis, scheduling, process control, robotics, and electronic circuit design. Such techniques can be used for the design of computational resources in a way that offers substantial promise for application to computing in nonlinear media, since the algorithms are almost independent of the medium in which the computation occurs. This is important in order to achieve effective computing in nonlinear media, since the algorithms do not need to directly manipulate the material to facilitate learning, and the task itself can be defined in a fairly unsupervised manner. In contrast, most traditional learning algorithms use techniques that require detailed knowledge of and control over the computing substrate involved. In this article we control the BZ and neuronal networks via a reinforcement learning approach, which uses evolutionary

computing to create generalizations over the state-action space—Holland's learning classifier system [22], in particular a form known as XCS [53].

The article is arranged as follows: The next section describes the subexcitable BZ system that forms the basis of the chemical computing aspect of our research. The next section describes a computational model of the system and XCS. Initial results from using the evolutionary learner to control the simulated and the real chemical systems are then presented. The form of neuronal cell culture we are developing is then described, along with the XCS control scenario used. Initial results from controlling the electrical stimulation via XCS are then presented, together with confirmation via another learning protocol presented in the literature. Finally, all results are discussed.

2 Chemical Experimental System

2.1 Materials and Equipment

Sodium bromate, sodium bromide, malonic acid, sulfuric acid, tris(bipyridyl) ruthenium(II) chloride, and 27% sodium silicate solution stabilized in 4.9 M sodium hydroxide were purchased from Aldrich (UK) and used as received unless stated otherwise.

An InFocus model projector was used to illuminate the computer-controlled image. Images were captured using a Panasonic NV-GS11 digital video camera. The microscope slide was immersed in the continuously-fed reaction solution contained in a custom-designed petri dish, designed by Radleys (Bristol, UK), with a water jacket thermostatted at 22°C. A Watson Marlow 205U multi-channel peristaltic pump was used to pump the reaction solution into the reactor and remove the effluent.

2.2 Experimental Procedures

2.2.1 Making Gels

A stock solution of the sodium silicate solution was prepared by mixing 222 mL of the purchased sodium silicate solution with 57 mL of 2 M sulfuric acid and 187 mL of deionized water, as in the procedure used by Wang et al. [51]. Precured solutions for making gels were prepared by mixing 5 mL of the acidified silicate solution with a solution consisting of 1.3 mL of 1.0 M sulfuric acid and 1.2 mL of 0.025 M tris(bipyridyl) ruthenium(II) chloride. Using capillary action, portions of this solution were transferred onto microscope slides with 100-μm shims and Plexiglas covers. The transferred solutions were left for 3 h to permit complete gellation, after which the covers and shims were removed and the gels washed in deionized water to remove residual tris(bipyridyl) ruthenium(II) chloride and the sodium chloride byproduct. The gels were 26 by 26 mm, with a wet thickness of approximately 100 μm. The gels were stored under water and rinsed right before use.

2.2.2 Catalyst-Free Reaction Mixture

The bulk of the catalyst-free reaction mixture was freshly prepared in 300-mL batches, which involved the in situ synthesis of stoichiometric bromomalonic acid from malonic acid and bromine generated from the partial reduction of sodium bromate. The catalyst-free reaction solution consisted of the 0.36 M sodium bromate, 0.0825 M malonic acid, 0.18 M sulfuric acid, and 0.165 M bromomalonic acid. To minimize deterioration during the experiment, this solution was kept in an ice bath. This solution was continuously fed into the thermostatted reactor, with a reactor residence time of 30 min.

2.2.3 Experimental Setup

The spatially distributed excitable field on the surface of the gel was made possible by the projection of a 10-by-10-cell checkerboard grid pattern generated using a computer. After Wang et al., the checkerboard image was composed of a heterogeneous network of cells at two light levels, a low intensity of 0.394 mW cm^{-2} and a high intensity of 9.97 mW cm^{-2}, representing excitable and non-excitable

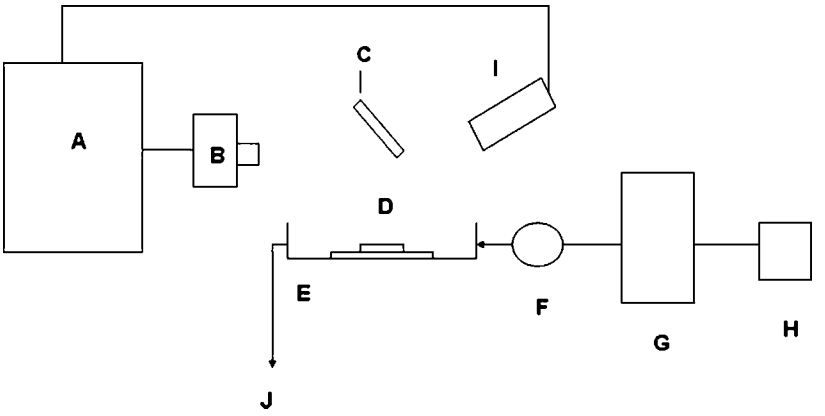

Figure 1. A block diagram of the experimental setup where the computer, projector, mirror, microscope slide with the catalyst-laden gel, thermostatted petri dish, peristaltic pump, thermostatted water bath, reservoir of catalyst-free reaction solution, digital camcorder, and effluent flow are designated by A, B, C, D, E, F. G, H, I, and J, respectively. The catalyst-free reaction solution reservoir was kept in an ice bath during the experiment.

domains, respectively. A digital video camera was used to capture the chemical wave fragments. A diagrammatic representation of the experimental setup is shown in Figure 1.

2.2.4 Data Capturing and Image Processing

A checkerboard grid pattern was projected onto the catalyst-laden gel through a 455-nm narrow bandpass interference filter and 100–100-mm-focal-length lens pair and mirror assembly. The projected grid was approximately 20 mm square. Every 10 s, the checkerboard pattern was replaced with a uniform gray level of 9.97 mW cm^{-2} for 400 ms, during which time an image of the BZ waves on the gel was captured. The purpose of removing the grid pattern during this period was to allow activity on the gel to be more visible to the camera and assist in subsequent image processing of chemical activity.

Captured images were processed to identify chemical wave activity. This was done by differencing successive images pixel by pixel to create a black-and-white thresholded image. Each pixel in the black-and-white image was set to white, corresponding to chemical activity, if the intensity of the red or blue channels differed in successive images by more than 1.95%. Pixels at locations not meeting this criterion were set to black. The thresholded images were automatically despeckled and manually edited to remove artifacts of the experiment, such as glare from the projector and bubbles from the oxidative decarboxylation of malonic acid and bromomalonic acid. The images were cropped to the grid location and the grid superimposed on the thresholded images to aid analysis of the results.

3 Computational System

3.1 Model

The features of this system were simulated using a two-variable Oregonator model modified to take account of the photochemistry [15, 29]:

$$\frac{\partial u}{\partial t} = \frac{1}{\varepsilon}\left(u - u^2 - (fv + \Phi)\frac{u - q}{u + q}\right) + D_u \nabla^2 u,$$

$$\frac{\partial v}{\partial t} = u - v.$$

The variables u and v represent the instantaneous local concentrations of the bromous acid autocatalyst and the oxidized form of the catalyst—$HBrO_2$ and tris(bipyridyl) Ru(III), respectively—scaled to dimensionless quantities. The ratio of the time scales of the two variables u and v is denoted by ε and depends on the rate constants and reagent concentration; f is a stoichiometric coefficient. The rate of the photoinduced bromide production is designated by Φ, which also denotes the excitability of the system, in which low light intensities facilitate excitation while high intensities result in the production of bromide that inhibits the process, as is experimentally verified. The scaling parameter q depends on reaction rates only. The system was integrated using the Euler method with a five-node Laplacian operator, time step $\Delta t = 0.001$, and grid point spacing $\Delta x = 0.15$. The diffusion coefficient D_u of species u was unity, while that of species v was set to zero, as the catalyst was immobilized in gel.

3.2 Evolutionary Algorithm

XCS represents a significant development of Holland's learning classifier system formalism and has been shown to be able to tackle many complex tasks effectively (see [6] for examples). It consists of a limited-size 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 ternary 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 selected either deterministically or stochastically, based on the fitness-weighted predictions (usually probability 0.5 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:

1. 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.
2. Rule predictions are then updated: $p_j = p_j + \beta(P - p_j)$
3. Each rule's accuracy κ_j is determined: $\kappa_j = \alpha(\varepsilon_0/\varepsilon)^{\nu}$, or $\kappa = 1$ where $\varepsilon < \varepsilon_0$. Here ν, α, and ε_0 are constants controlling the shape of the accuracy function.
4. A relative accuracy κ_j' is determined for each rule by dividing its accuracy by the total of the accuracies in the action set.
5. The relative accuracy is then used to adjust the classifier's fitness F_j using the *moyenne adaptive modifiée* (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 inversely proportional to the error in reward prediction, with errors below ε_0 not improving 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 [52] in its reinforcement procedure.

A genetic algorithm (GA) [21] acts in action sets [A], that is, 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 (after [5]), based on the average time since the members of the niche last participated in a GA.

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 as neural networks—see [45] for an overview). The reader is referred to [8] for an algorithmic description of XCS, and [7] for an overview of current formal understanding of its operations.

3.3 XCS Control: Simulator

The aforementioned model of the BZ system has been interfaced to an implementation of XCS in a way that approximates the envisaged hardware-wetware scenario. A 3 × 3 grid is initialized with a pulse of excitation in the bottom middle cell, as in the wetware experiments described in Section 2. Two light levels have been used thus far: one that is sufficiently high to inhibit the reaction, and the other low enough to enable it. The modeled chemical system is then simulated for 10 s of real time. A 9-bit binary description of the 3 × 3 grid is then passed to the XCS. Each bit corresponds to a cell, and it is set to "true" if the average level of activity within the given cell is greater than a predetermined threshold. The XCS returns a 9-bit action string, each bit of which indicates whether light of the high ($\Phi = 0.197932817$) or low ($\Phi = 0.007791641$) intensity should be projected onto the given cell. Another 10 s of real time is then simulated, and so on, until either a maximum number of iterations has passed or the emergent spatial-temporal dynamics of the system match a required configuration. In this initial work, a fragment is required to exist in the middle left-hand cell of the grid *only*. At such a time, a numerical reward of 1,000 is given the system, and the system is reset for another learning trial. To be able to obtain this behavior reliably, it has been found beneficial to use an intermediate reward of 500 in the presence of a fragment in the target cell, regardless of the activity on the rest of the grid (see [14] for related discussions). The XCS parameters used for this were (largely based on [53]): $N = 30{,}000$, $\beta = 0.2$, $\mu = 0.04$, $\chi = 0.8$, $\gamma = 0.71$, $\theta_{del} = 20$, $\delta = 0.1$, $\varepsilon_0 = 10$, $\alpha = 0.1$, $\nu = 5.0$, $\theta_{mna} = 512$, $\theta_{GA} = 25$, $\rho_I = \varepsilon_I = F_I = 10.0$, $p_{\#} = 0.33$. Other parameters for the BZ model were $\varepsilon = 0.022$, $f = 1.4$, $q = 0.002$.

Figure 2 shows a typical light program and associated wave fragment behavior sequence, here taking nine steps to solve the problem, which appears optimal with the given parameters of the simulator and allowed time between XCS control steps. Figure 3 shows the average reward received by the learner per trial on exploit trials only, as a 50-point moving average (after [53]) over three runs. As can be seen, this approaches the maximum of 1,000 in the time considered; the XCS controller is reliable in its ability to develop a fragment controller in the given scenario.

3.4 XCS Control: Chemistry

Given the success of the simulation experiments, the XCS was connected to the chemical system described in Section 2. The scenario for the chemical experiments was the same as for the simulations, although it must be noted that there is a slightly longer delay before the XCS is able to control the initial light levels once the pulse is added to the grid, due to the image processing required. Figure 4 shows an example result using the same parameters as before. This solution was discovered on trial 4 and then refined over the subsequent three trials to that shown. It can be seen that XCS has learned to control the fragment in the real chemical system as it did in the simulation.

4 Neuronal System

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, that is, cells in a network grown across the surface of a multi-electrode

array dish. However, it has long been known that aggregated (i.e., 3D) neuronal cell cultures exhibit properties that are remarkably similar to their in vivo counterparts. For example, early studies showed structures identical to hippocampal architecture [11], and Seeds [38] 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., [34, 47]).

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 behavior in aggregated cultures of such cells is typically very similar to that reported in monolayers of mammalian cortical

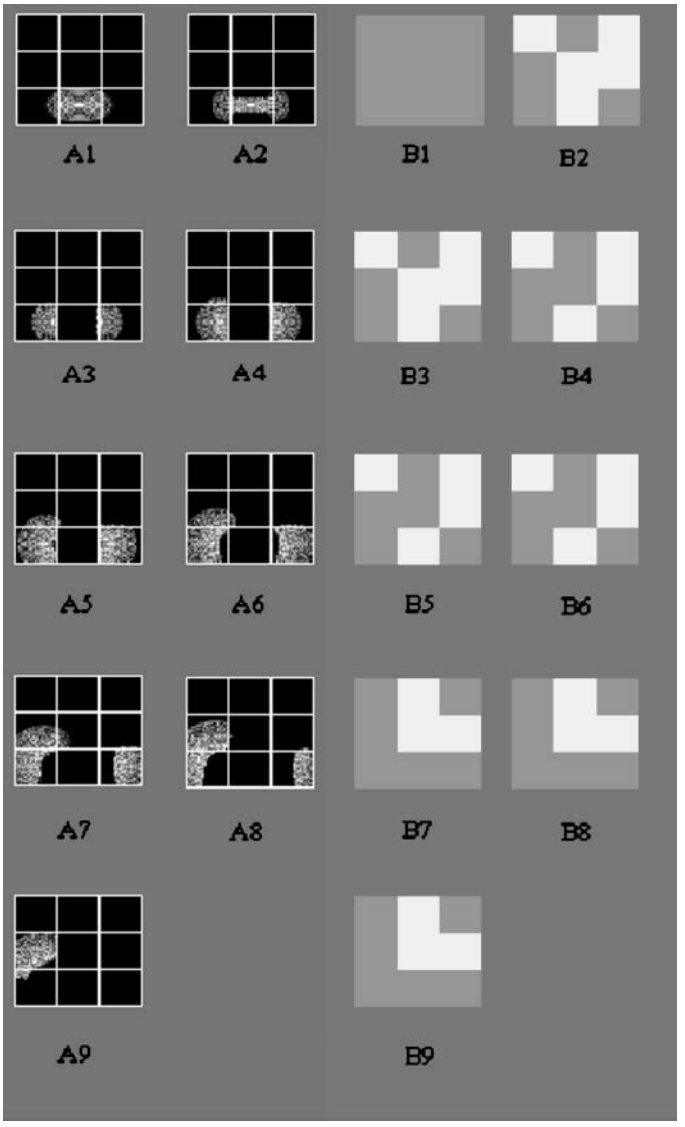

Figure 2. Example control of the simulated chemical system (A) under the learned light programs (B).

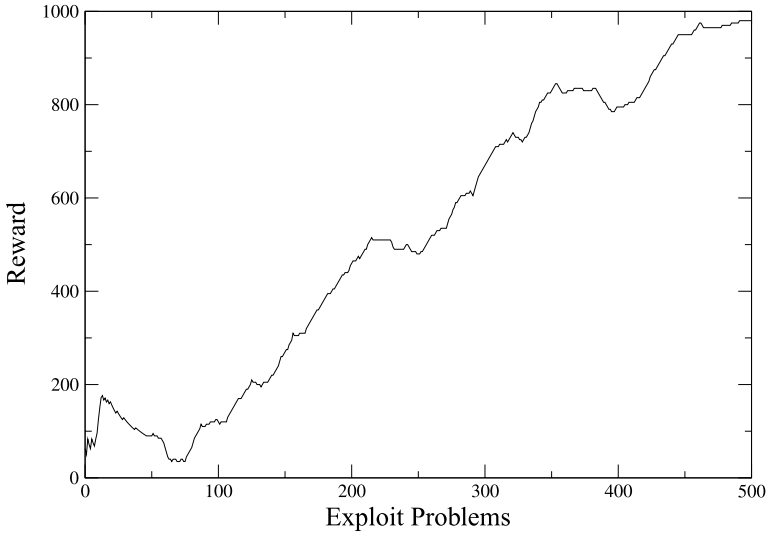

Figure 3. The average reward received by the XCS controllers with increasing number of learning cycles or problems.

cells [49]. However, response to simple stimulation has been shown to be capable of causing 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 5 shows an example of the aggregates used in this study.

4.1 Materials and Equipment

Fertilized eggs of *Gallus domesticus* (Red Island hens) were obtained from a local poultry farm and transported to our laboratories; eggs underwent a preconditioning phase at room temperature. After marking and recording, the eggs were incubated for 7 days at 37°C in an egg incubator (Octagon 100, Brinsea Ltd., UK). During the incubation, appropriate levels of humidity were maintained.

On the day of spheroid preparation, eggs were removed from the incubator at a precise time in order that the E7 embryonic stage had been reached. The embryo was removed using aseptic techniques and placed in a 100-mm petri dish with Hanks balanced salt solution (HBSS, Life Technologies). Several embryos were pooled and then washed three times in HBSS. Under sterile conditions, the neuroepithelial tissue was removed under a stereo Leica Zoom 2000 microscope (Leica Co., Germany) with transmission and reflected illumination. The tissue was cleaned of meninges and transferred to a 25-mm petri dish with culture medium [DMEM, nutrient mixture F-12 Ham, fetal bovine serum, L-glutamine solution, penicillin-streptomycin solution, progesterone (water soluble), putrescine dihydrochloride, 3,3,5-triiod-L-thyronine, selenium dioxide, holotransferrin (human), and insulin, all from SIGMA]. The tissue was then collected and placed in a test tube with 5 mL of culture medium and triturated to prepare a cell suspension that was then filtered through a 35-μm Nybolt membrane. A trypan blue dye (0.4% w/v) exclusion assay was conducted and a cell count performed under light microscopy using a hemacytometer. The plating cell concentration was 0.5×10^6 viable cells per milliliter of cell culture medium, 3 mL per well in a six-well culture plate. The six-well plates were placed on a gyrorotatory shaker (Innova 2000, New Brunswick Scientific Co., Inc.) and cultured at 37°C in 5% CO_2 and 95% air at 75 rpm. The cell culture medium was refreshed every other day by removing some 50% of the old medium and refilling with fresh medium warmed to a physiological temperature.

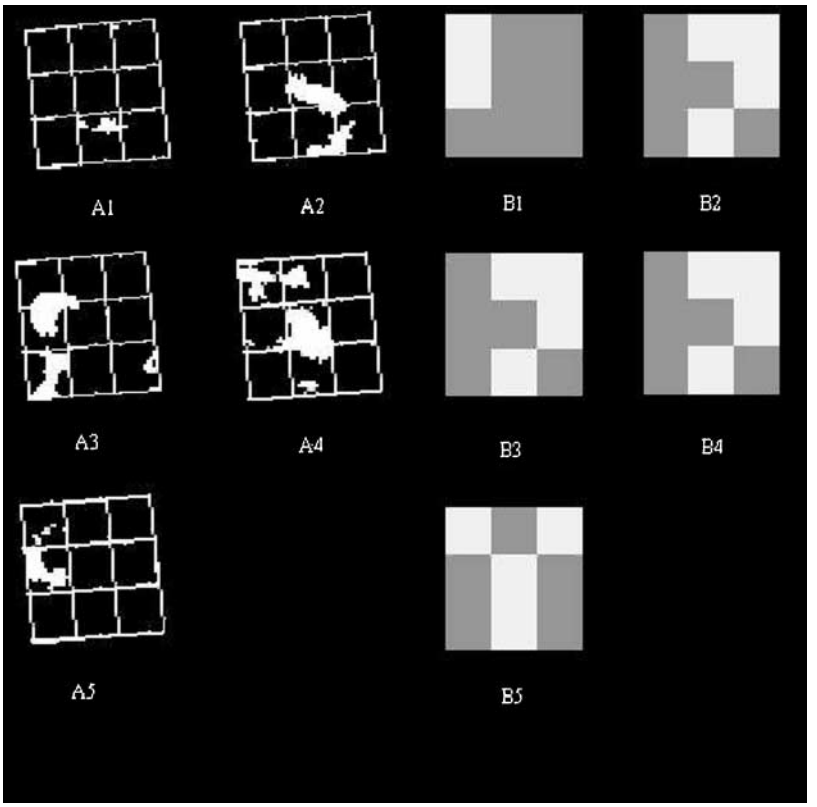

Figure 4. Example control of the real chemical system (A) under the learned light programs (B). Skewing of cell images is due to camera setup.

Multi-electrode arrays (MEAs) (MCS-2100, Multi Channel Systems MCS GmbH, Aspenhaustrasse 21, 72770 Reutlingen, Germany) with pyramidal electrodes (40 × 40 × 70 μm, spaced on 200 μm) were used to record electrical activity of the spheroids. The MEA dish surface was modified with 10-μg/mL aqueous solution of polymer ethylene imine (PEI) (Fluka Chemie AG, Buchs, Switzerland)

Figure 5. Scanning electron microscope image of DIV21 hen embryo aggregate neuronal culture.

under sterile conditions. The molecular weight of the 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 spheroids.

The electrical recording from the spheroids was performed with a 60-channel data acquisition system, where the sampling frequency of each channel was set to 25 kHz and the single-channel amplification kept at 1,200 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 300-Hz filter was performed in a way that was similar to other studies (e.g., [26]). The spikes were detected by a threshold depending upon the standard deviation and the offset of noise. A set of data was monitored and raw signal, filtered signal, and spikes were 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 analysis. The recorded signals were analyzed and the spike parameters were extracted using the MC Rack software and analyzed with bespoke software as a postprocessing step.

The stimulation of the hen embryo brain spheroids was realized with the eight-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 stimulation protocol shared features of relevant published studies (e.g., [50]), consisting of a single sharp biphasic impulse of 300-μs duration and voltages between 300 and 2,000 mV for each phase per sweep of 1 s.

4.2 XCS Control: Scheme

In the current study XCS was 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-s window. Typically, an individual aggregate covers three or four electrodes in a dish, as shown in Figure 6, one or two of which will show a suitably good connection into the neuronal network therein, that is, spikes will be detected of the kind shown in Figure 7. 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 equal to the spontaneous mean plus two standard deviations; a significant increase in typical spiking frequency is required under stimulation.

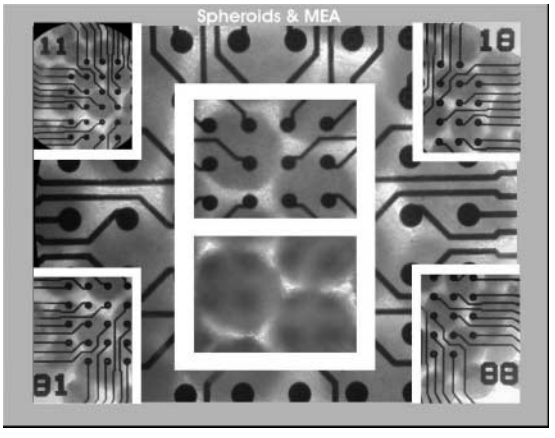

Figure 6. Composite picture of phase contrast microscopy images taken at various optical magnifications and focal planes of aggregate cell cultures on multi-electrode array dish.

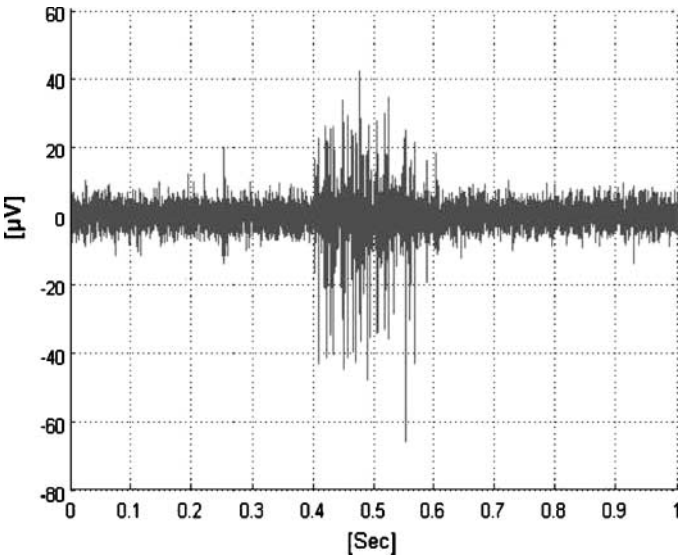

Figure 7. Example spiking behavior recorded on a single electrode.

The input to the XCS on each cycle is the spiking frequency of the neuronal network averaged over the last 3 s 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 s of spontaneous behavior, and the second as a fraction of the maximum allowed stimulation time of 600 s. 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 in the last stimulation period over that immediately prior, and a reward of 1,000 is given if the target spiking frequency, or a greater one, was achieved.

Following [39], we allow a 300-s rest period between applications of the stimulus and truncate the maximum duration of stimulation to 600 s. Thus 300 s 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 s 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 unordered pairs of real numbers in the range [0, 1], one pair for each environmental input (after [43]). A pair is considered to match the corresponding input value if one member of the pair is smaller than or equal to the target, and the other is larger or equal. The action of the classifier is an integer. Given the online nature of the task, roulette wheel rather than random action selection was used in explore trials.

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 centered on the current value, or a fixed small change (here, 0.1). Action mutation is done 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 center, and two values are created in the range $[e - C_{\max}, e + C_{\max}]$, where $C_{\max}$ is 0.1.

The XCS parameters used were again typical for those in the literature: $N = 3{,}000$, $\beta = 0.2$, $\mu = 0.04$, $\chi = 0.8$, $\theta_{del} = 20$, $\delta = 0.1$, $\varepsilon_0 = 10$, $\alpha = 0.1$, $\nu = 5.0$, $\theta_{mna} = 3$, $\theta_{GA} = 2$, $\rho_I = \varepsilon_I = F_I = 10.0$. The cell cultures used were all in the range of 20 to 30 days old in vitro.

4.3 XCS Control: Results

After a number of experiments it became clear that the XCS was only able to alter a neuronal network's behavior in roughly a third of cases. Figure 8 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 over which the stimulation is applied. However, Figure 9 shows a case where no significant change in spiking appears to have occurred, regardless of how the XCS adjusted the network's stimulation. In other cases the average spiking response decreases during the experiment regardless of the stimulation duration (not shown). These figures show both the exploration and exploitation trial behavior of XCS, that is, the actual online

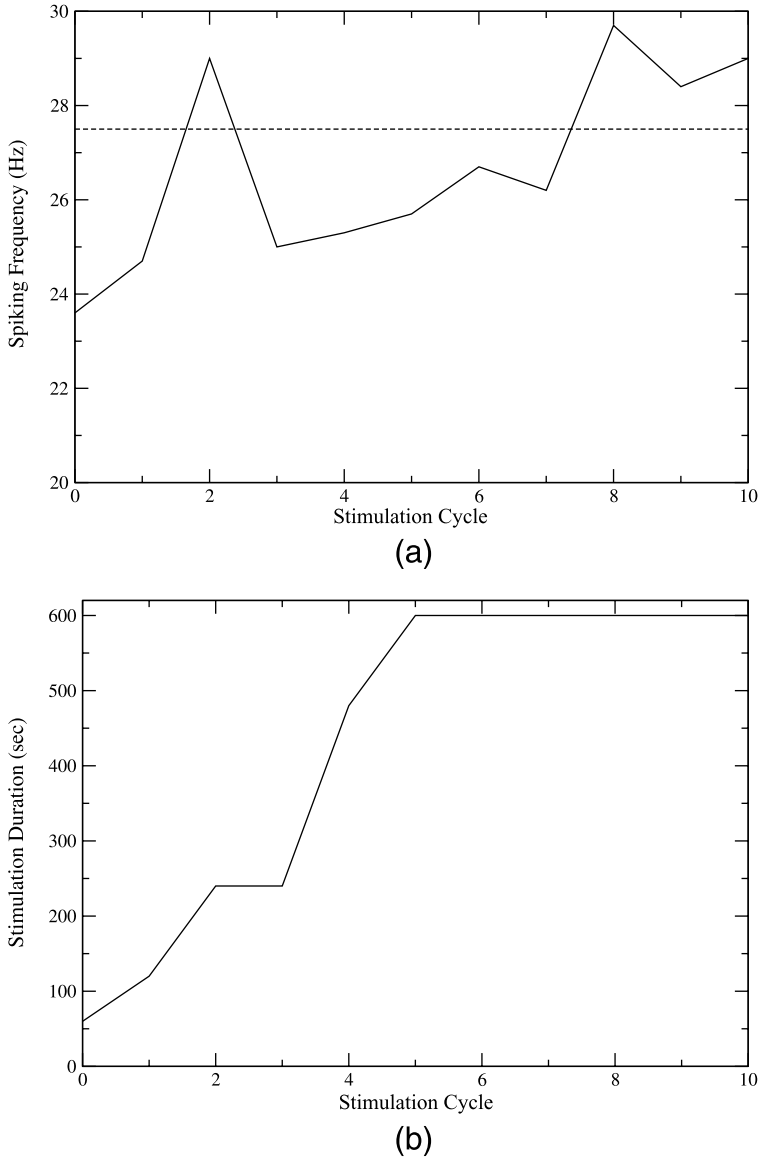

Figure 8. Example learning behavior under XCS control, showing (a) the spiking-frequency response becoming repeatedly higher than the target indicated by the dashed line, and (b) how XCS altered the stimulus application time to achieve this.

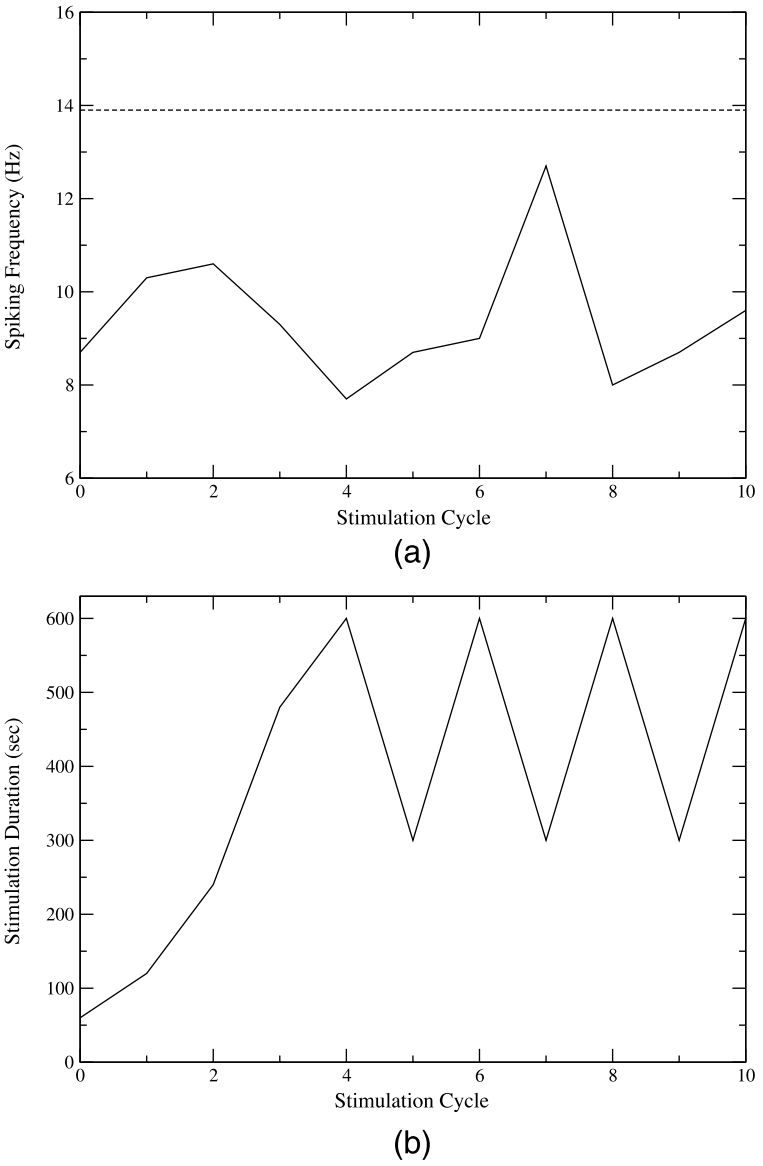

Figure 9. Example unsuccessful learning behavior under XCS control, showing (a) the spiking-frequency response never rising to the target indicated by the dashed line, and (b) how XCS altered the stimulus application time.

stimulation as experienced by the neuronal networks; unlike the chemical systems above, every stimulation has a potential long-term effect on the neuronal network.

Given these findings, we implemented the aforementioned behavior shaping and learning protocol of Shahaf and Marom [39] for comparison. This scheme, in contrast to the more widespread consideration of neuromodulatory reward mechanisms for learning, is inspired by the work of behaviorists such as Clark Hull (e.g., [24]). Known as the stimulus regulation principle (SRP), it proposes that reward, and hence learned, behavior are achieved through the removal of the driving stimulus. That is, neurons cease a continual alteration of their connectivity when the driving stimulus is removed, and hence the behavior 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 behavior for 300 s. 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 s, as before) or if 600 s had elapsed. Again, a 300-s rest period between applications was allowed.

Figure 10 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 that target 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

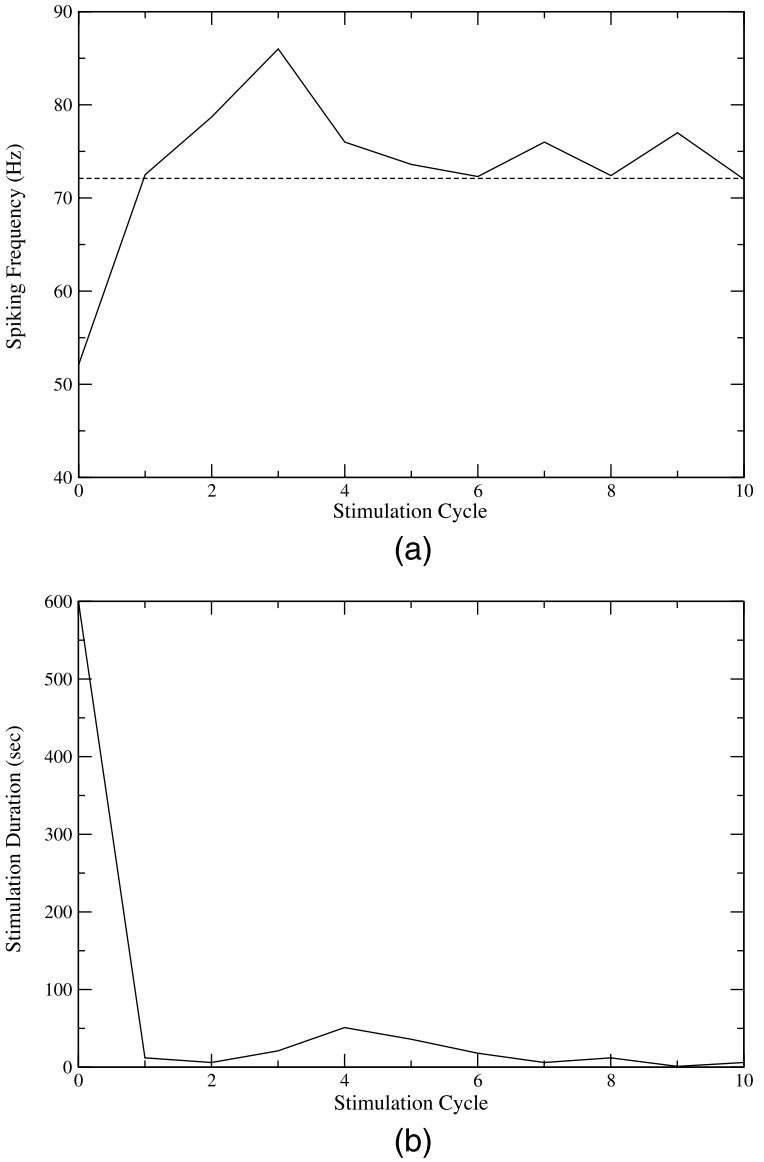

Figure 10. Example learning behavior under SRP control, showing (a) the spiking-frequency response becoming repeatedly higher than the target indicated by the dashed line and (b) how the stimulus was applied and removed to achieve this. The spiking frequency shown is the last recorded on a given cycle.

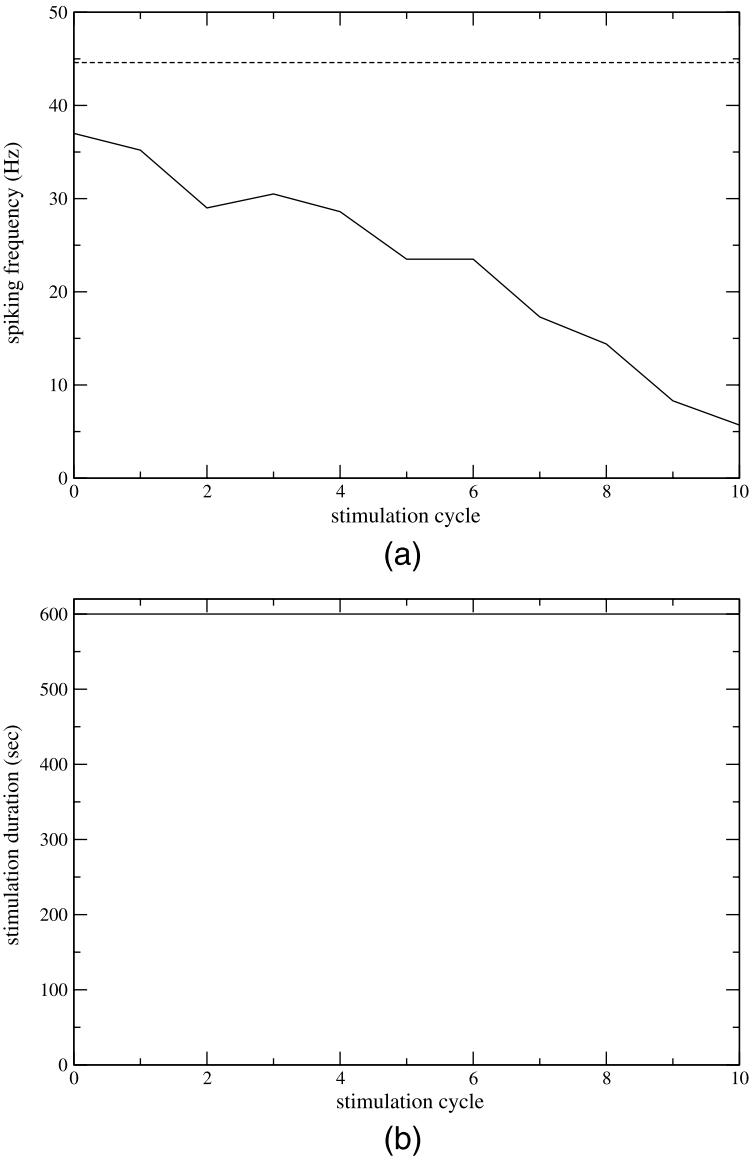

Figure 11. Example unsuccessful learning behavior under SRP control, showing (a) the spiking-frequency response falling away from the target indicated by the dashed line, and (b) how the stimulus was therefore constantly applied throughout. The spiking frequency shown is the last recorded on a given cycle.

again found that such results occurred only about a third of the time. Figure 11 shows an example where the target frequency is never seen: The stimulation always remains applied for the maximum of 600 s, and the spiking frequency drops over time. Examples with no significant change were again also seen (not shown).

5 Conclusions

Excitable and oscillating chemical systems have previously been used to perform a number of very simple computational tasks. We propose that utilizing networks of coupled oscillating chemical reactions will open chemical computing to wider domains and are interested in using light to construct such

systems on the surface of a gel—cells in a regular grid here. In this article we have presented initial results from a methodology by which to achieve the complex task of designing such systems—through the use of evolutionary learning techniques. We have shown, using both simulated and real systems, that it is possible to control the behavior of a light-sensitive BZ reaction using XCS.

Neuronal networks represent another form of unconventional system by which to achieve computation. However, we suggest that the use of two-dimensional networks is a somewhat impoverished formalism and that three-dimensional networks represent a potentially fruitful avenue of research. The results from using XCS again, and then an SRP-inspired protocol, to induce learning indicate three possible rudimentary responses to simple stimulation from such in vitro neuronal networks: *excitation*, giving the potential for significant increases in typical spiking behavior; *inhibition*, wherein spiking behavior decreases due to stimulation; and *unchanging*, that is, no significant shift in spiking behavior over spontaneous behavior is seen due to the stimulus. We found that each such behavior was equally likely and that neither learning protocol could affect the underlying behavior of a given neuronal network.

Jimbo et al. [25], using monolayer cultures of mammalian cortex, reported simple stimulation at an electrode could induce either an excitatory or an inhibitory response at other electrodes—"pathway-dependent plasticity." Our results therefore suggest that the same is also true for aggregate neuronal cell cultures, but that a third class of behavior is possible. It is not clear at this time whether such behavior is typical in monolayers.

We are currently exploring the use of XCS to elicit more subtle responses to stimulation from in vitro neuronal networks and to demonstrate computation; we have recently demonstrated rudimentary computational capability in neuronal networks using the simple SRP (see Appendix). We are also extending the chemical scenario presented to include larger grids containing many more concurrent fragments of excitation. In the longer term, our aim is to study information processing in nonlinear chemical and neuronal media in order to determine fundamental principles for the design of novel computers based on nonlinear media and to abstract new unconventional computing mechanisms.

Acknowledgment

This work was supported by EPSRC grant no. GR/T11029/01.

References

1. Adamatzky, A. (2001). *Computing in nonlinear media and automata collectives*. Bristol, UK: IoP Publishing.
2. Agladze, K., Aliev, R. R., Yamaguhi, T., & Yoshikawa, K. (1996). Chemical diode. *Journal of Physical Chemistry*, *100*, 13895–13897.
3. Bar-Eli, K., & Reuveni, S. (1985). Stable stationary-states of coupled chemical oscillators: Experimental evidence. *Journal of Physical Chemistry*, *89*, 1329–1330.
4. Bar-Eli, K. (1985). On the stability of coupled oscillators. *Physica D*, *14*, 242–252.
5. Booker, L. (1989). Triggered rule discovery in classifier systems. In R. Schaffer (Ed.), *Proceedings of the International Conference on Genetic Algorithms* (pp. 265–274). San Mateo, CA: Morgan Kaufmann.
6. Bull, L. (Ed.) (2004). *Applications of learning classifier systems*. Berlin: Springer.
7. Bull, L., & Kovacs, T. (Eds.) (2005). *Foundations of learning classifier systems*. Berlin: Springer.
8. Butz, M., & Wilson, S. W. (2002). An algorithmic description of XCS. *Soft Computing*, *6*(3–4), 144–153.
9. Crowley, M. F., & Field, R. J. (1986). Electrically coupled Belousov-Zhabotinskii oscillators 1: Experiments and simulations. *Journal of Physical Chemistry*, *90*, 1907–1915.
10. Crowley, M. F., & Epstein, I. R. (1989). Experimental and theoretical studies of a coupled chemical oscillator: Phase death, multistability and in-phase and out of phase entrainment. *Journal of Physical Chemistry*, *93*, 2496–2502.
11. DeLong, G. R. (1970). Histogenesis of fetal mouse isocortex and hippocampus in reaggregating cell cultures. *Developmental Biology*, *22*, 563–583.

12. 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.

13. Dolnik, M., & Epstein, I. R. (1996). Coupled chaotic oscillators. *Physical Review E*, *54*, 3361–3368.

14. Dorigo, M., & Colombetti, M. (1998). *Robot shaping*. Cambridge, MA: MIT Press.

15. Field, R. J., & Noyes, R. M. (1974). Oscillations in chemical systems. IV: Limit cycle behavior in a model of a real chemical reaction. *Journal of Physical Chemistry*, *60*, 1877–1884.

16. Gorecki, J., Yoshikawa, K., & Igarashi, Y. (2003). On chemical reactors that can count. *Journal of Physical Chemistry A*, *107*, 1664–1669.

17. Hjelmfelt, A., & Ross, J. (1993). Mass-coupled chemical systems with computational properties. *Journal of Physical Chemistry*, *97*, 7988–7992.

18. Hjelmfelt, A., Weinberger, E. D., & Ross, J. (1991). Chemical implementation of neural networks and Turing machines. *Proceedings of the National Academy of Sciences of the U.S.A.*, *88*, 10983–10987.

19. Hjelmfelt, A., Weinberger, E. D., & Ross, J. (1992). Chemical implementation of finite-state machines. *Proceedings of the National Academy of Sciences of the U.S.A.*, *89*, 383–387.

20. Hjelmfelt, A., Schneider, F. W., & Ross, J. (1993). Pattern-recognition in coupled chemical kinetic systems. *Science*, *260*, 335–337.

21. Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor: University of Michigan Press.

22. Holland, J. H. (1986). Escaping brittleness: The possibilities of general-purpose learning algorithms applied to parallel rule-based systems. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning, an artificial intelligence approach* (pp. 48–78). San Mateo, CA: Morgan Kaufmann.

23. Holz, R., & Schneider, F. W. (1993). Control of dynamic states with time-delay between 2 mutually flow-rate coupled reactors. *Journal of Physical Chemistry*, *97*, 12239.

24. Hull, C. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.

25. 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.

26. Jimbo, Y., Kawana, A., Parodi, P., & Torre, V. (2000). The dynamics of a neuronal culture of dissociated cortical neurons of neonatal rats. *Biological Cybernetics*, *83*, 1–20.

27. Jung, P., Cornell-Bell, A., Madden, K. S., & Moss, F. J. (1998). Noise-induced spiral waves in astrocyte syncytia show evidence of self-organized criticality. *Neurophysiology*, *79*, 1098–1101.

28. Kawato, M., & Suzuki, R. (1980). Two coupled neural oscillators as a model of the circadian pacemaker. *Journal of Theoretical Biology*, *86*, 547–575.

29. Krug, H.-J., Pohlmann, L., & Kuhnert, L. (1990). Analysis of the modified complete Oregonator accounting for oxygen sensitivity and photosensitivity of Belousov-Zhabotinsky systems. *Journal of Physical Chemistry*, *94*, 4862–4866.

30. Kuhnert, L., Agladze, K. I., & Krinsky, V. I. (1989). Image processing using light sensitive chemical waves. *Nature*, *337*, 244–247.

31. Laplante, J. P., Pemberton, M., Hjelmfelt, A., & Ross, J. (1995). Experiments on pattern recognition by chemical kinetics. *Journal of Physical Chemistry*, *99*, 10063–10065.

32. Lebender, D., & Schneider, F. W. (1994). Logical gates using a nonlinear chemical reaction. *Journal of Physical Chemistry*, *98*, 7533–7537.

33. Michalewicz, Z., & Fogel, D. (1999). *How to solve it*. Berlin: Springer-Verlag.

34. Moscona, A. (1961). Rotation mediated histogenic aggregation of dissociated cells. *Experimental Cell Research*, *22*, 455–475.

35. Motoike, I. N., Yoshikawa, K., Iguchi, Y., & Nakata, S. (2001). Real time memory on an excitable field. *Physical Review E*, *63*(036220), 1–4.

36. Rossler, O. E. (1974). Chemical automata in homogeneous and reaction-diffusion kinetics. In M. Conrad, W. Guttinger, & M. Dal Cin (Eds.), *Physics and mathematics of the nervous system* (pp. 399–418, 546–582). Berlin: Springer.

37. Ruaro, M. E., Bonifazi, P., & Torre, V. (2005). Toward the neurocomputer: Image processing and pattern recognition with neuronal cultures. *IEEE Transactions on Biomedical Engineering*, *52*(3), 371–383.

38. Seeds, N. W. (1971). Biochemical differentiation in reaggregating brain cell culture. *Proceedings of the National Academy of Sciences of the U.S.A.*, *68*(8), 1858–1861.

39. Shahaf, G., & Marom, S. (2001). Learning in networks of cortical neurons. *Journal of Neuroscience*, *21*(22), 8782–8788.

40. Sielewiesiuk, J., & Gorecki, J. (2002). Passive barrier as a transformer of chemical frequency. *Journal of Physical Chemistry A*, *106*, 4068–4076.

41. Steinbock, O., Toth, A., & Showalter, K. (1995). Navigating complex labyrinths: Optimal paths from chemical waves. *Science*, *267*, 868–871.

42. Steinbock, O., Kettunen, P., & Showalter, K. (1996). Chemical wave logic gates. *Journal of Physical Chemistry*, *100*, 18970–18975.

43. Stone, C., & Bull, L. (2003). For real! XCS with continuous-valued inputs. *Evolutionary Computation*, *11*(3), 299–336.

44. Stuchl, I., & Marek, M. (1982). Dissipative structures in coupled cells: Experiments. *Journal of Physical Chemistry*, *77*, 2956–2963.

45. Sutton, R., & Barto, A. (1998). *Reinforcement learning.* Cambridge, MA: MIT Press.

46. Takayama, Y., & Jimbo, Y. (2006). Modification of evoked responses induced by correlated stimuli in cultured cortical networks. *Proceedings of the 5th International Meeting on Substrate-Integrated Micro Electrode Arrays* (pp. 22–25).

47. 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.

48. Toth, A., Gaspar, V., & Showlater, K. (1994). Signal transmission in chemical systems: Propagation of chemical waves through capillary tubes. *Journal of Physical Chemistry*, *98*, 522–531.

49. 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.

50. 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.

51. Wang, J., Kadar, S., Jung, P., & Showalter, K. (1999). Noise driven avalanche behavior in subexcitable media. *Physical Review Letters*, *82*, 855–858.

52. Watkins, C. J. (1989). *Learning from delayed rewards.* Ph.D. Thesis. Cambridge University, Cambridge, UK.

53. Wilson, S. W. (1995). Classifier fitness based on accuracy. *Evolutionary Computing*, *3*, 149–175.

54. Zaikin, A. N., & Zhabotinsky, A. M. (1970). Concentration wave propagation in two-dimensional liquid-phase self-oscillating system. *Nature*, *225*, 535–537.

Appendix

Using a doubled version of the stimulation protocol described in Section 4.1, we have been able to use the SRP to train (excitable) networks to perform a two-input Boolean OR function. Double stimulation is given concurrently at the electrodes under the identified aggregate for a logical 11 input, and single stimulation for 01 and 10. The input pattern cycles through these three in each training period. Input 00 is taken as the unstimulated state. Figure 12 shows an example result wherein, after training, a spiking response over the threshold is almost immediately obtained for any of the three inputs 11, 01, and 10, all of which would require a response representing a logical 1 for logical OR. Hence if a spiking response over the threshold of the mean plus two standard deviations, as before, is deemed to represent a logical "1" and anything else, logical "0," the network can be seen to perform OR.

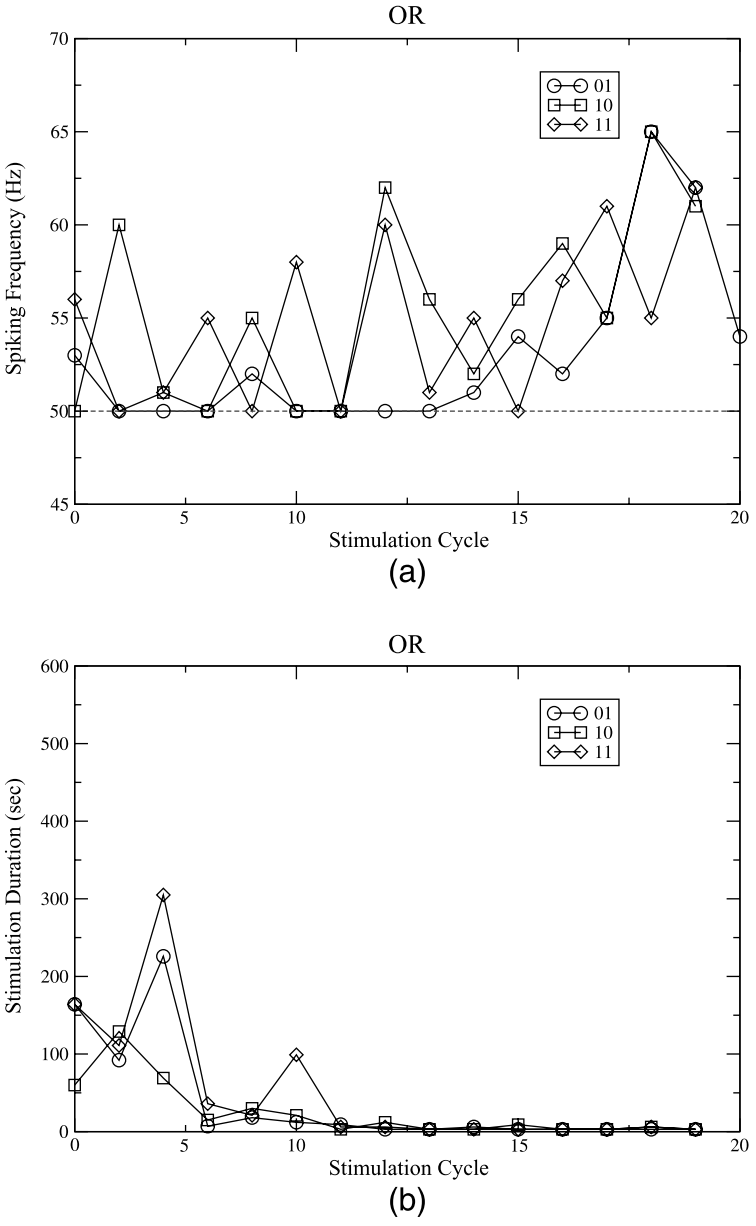

Figure 12. Example successful learning of a two-input OR function using the SRP, showing (a) that the spiking-frequency response is repeatedly higher than the target indicated by the dashed line for all three inputs containing stimulation (i.e., logical "1"), and (b) how the stimulus was applied and removed to achieve this. The spiking frequency shown is the last recorded on a given cycle.