

# Steering the Growth of Adaptive Self-Preserving Dissipative Structures

Matthew Egbert<sup>1,2,\*</sup>, Yan Kolezhitskiy<sup>1</sup> and Nathaniel Virgo<sup>3</sup>

<sup>1</sup> School of Computer Science, University of Auckland, Auckland, NZ

<sup>2</sup>Te Ao Mārama — Centre for Fundamental Inquiry, University of Auckland, Auckland, NZ

<sup>3</sup> Earth-Life Science Institute, Tokyo Institute of Technology, Tokyo, Japan

\* m.egbert@auckland.ac.nz

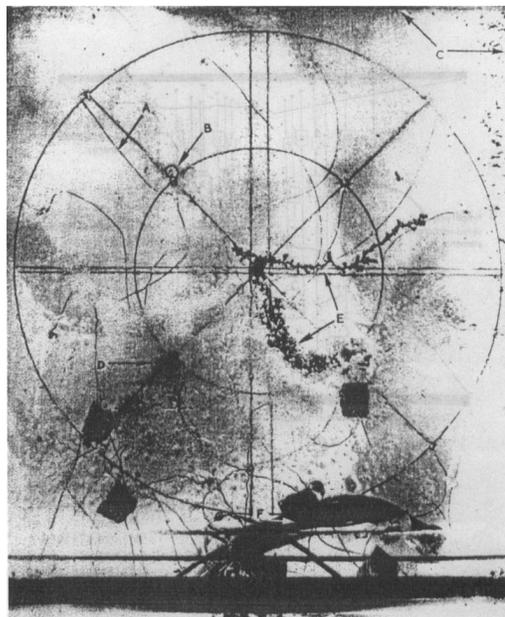
## Abstract

In the 1950s, the famous cyberneticists Gordon Pask and Stafford Beer conducted a series of remarkable electrochemical deposition experiments. By applying an electric potential across electrodes submerged in an acidic solution of ferrous sulfate, they could bias the growth of electrochemical deposition so as to form functional structures including sensory structures capable of distinguishing between different sounds. Unfortunately, the details of their apparatus and methods are unavailable. As a consequence, their experiment has not been replicated, and the precise mechanisms underlying their results remain unknown. As preliminary steps toward recreating their remarkable results, this paper presents a new computational model that simulates the growth and decay of dendritic structures similar to those investigated by Beer & Pask. We use this model to demonstrate a plausible mechanism through which an electrochemical system of this kind could respond to a reinforcement signal. More specifically, we investigate three strategies for varying the applied electrical current so as to guide the formation of structures into target forms. Each presented strategy succeeds at influencing the growth of the structure, with the most successful strategy involving a ‘constant-current’ feedback mechanism combined with an externally driven oscillation. In the discussion, we compare the adaptation of these structures with various biological adaptive processes, including evolution and metabolism-based adaptive behaviour.

## Introduction

The research presented below is inspired by the electrochemical deposition experiments undertaken by Gordon Pask and Stafford Beer in the 1950s. By applying a current across electrodes placed in an acidic solution of ferrous sulfate, these researchers induced the electrochemical deposition of iron onto the negatively charged electrode(s) (see Fig. 1). By varying the applied voltage, they could choose when iron was deposited and when it dissolved back into solution, and using this technique to reward (i. e., stabilize) desired growth and punish (i. e., dissolve) less desirable growth, they grew an iron ‘ear’ that was capable of distinguishing between two different frequencies of sound.

“We have made an ear and we have made a magnetic receptor. The ear can discriminate two frequencies, one



**Figure 1: Photograph of electrochemical deposition experiments conducted by Gordon Pask and Stafford Beer in the 1950s.** From (Pask, 1958) and (Cariani, 1993).

of the order of fifty cycles per second and the other of the order of one hundred cycles per second. The ‘training’ procedure takes approximately half a day and once having got the ability to recognize sound at all, the ability to recognize and discriminate two sounds comes more rapidly ... The ear, incidentally, looks rather like an ear. It is a gap in the thread structure in which you have fibrils which resonate at the excitation frequency.” (Pask, 1959, p. 261)

From their description, Pask and Beer’s system seems to exhibit a kind of reinforcement learning: it develops specific structures (such as fibrils with a particular resonant frequency) in response to a reward signal (current), without being given precise instructions on what form the structures should take. This is remarkable because it is a relatively sim-

ple physical system and lacks any obvious system for assigning credit or propagating the reward signal. Our goal is to understand what makes this possible, and whether the principles underlying it will generalise to other kinds of physical or dynamical system.

In biological evolution, a directed process of selection biases undirected (i. e., random) genetic variation, resulting in the formation of adaptive structures. Pask and Beer's experiments might be described in similar terms: iron accrues rather randomly, providing a kind of stochastic growth—and direction is given to this change via the selective variation of the applied current. This evolution-inspired description given above seems plausible, but growing an 'ear' merely by changing a voltage applied to a solution of ferrous sulfate seems to almost border on magic. How did this system work and what are the limits of this technique for growing functional structures, using only a reward/punishment-like feedback? If it is a kind of physical instantiation of a search algorithm, what is the search space like—are there lots of local minima or is it better captured as a very high-dimensional space full of neutral networks (Huynen, 1996) that facilitate adaptation to a wide variety of selection pressures? Does the evolutionary metaphor completely describe or explain how these kinds of systems might adapt to different 'selection pressures' (i. e., reward schemes), or are there other dynamics that don't fit so nicely into an evolutionary metaphor.

We do not yet know the answer these questions. Unfortunately, Beer and Pask failed to publish this area of their research in sufficient detail that others might repeat it. In the next section, we present our first steps toward recreating Beer and Pask's results, in the form of a computational model of electrochemical deposition. We use this model to evaluate the possibility of using dynamically modulated voltage so as to steer the growth of the structure in desired forms. But before we delve into the details of our model, we first review some related research which highlights some reasons that we find this system interesting.

## Related research

There have been a number of efforts to recreate Pask's work—see e. g. the projects listed in (Boden, 2010, p. 136). In the cases that we are familiar with, there has been success in growing dendritic structures, but nothing as remarkable as the development of a new sensor as reported by Pask and Beer. There have, on the other hand, been a number of interesting results in the investigation of a comparable adaptive self-organising dissipative structure (Nicolis and Prigogine, 1977) known as a ramified charge-transportation networks.

A ramified-charge-transportation-network (RCTN) consists of a number of small steel spheres, placed in a petri dish, and partially submerged in castor oil. A circular grounded electrode runs around the periphery of the dish and a high-voltage ( $\approx 20kV$ ) electrode sits above it. These simple systems demonstrate remarkable self-organizing dy-

namics. Specifically, the spheres self-organize into tree-like structures with topology that depends upon initial configuration of the beads, and that can be radically different based on minor changes in the initial setup (Jun and Hübler, 2005). Once grown these structures display statistically robust network features. For example, when the number of spheres is kept roughly the same, the number of termini and branching points will remain similar despite any topologically different structures (Jun and Hübler, 2005).

Kondepudi et al. (2015) describe the dynamics of these systems in terms of 'energy-seeking' and 'self-healing' behaviours. If a branch is broken, then the system will restore it. Further, the tree will continue moving its branches around the available space to maximize the current conducted by the structure. This can be considered a form of self-preservation, as the flow of electricity is what allows the system to persist in spite of its ordered, low-entropy state. From these and other observations, Kondepudi et al. argue that their overall behaviour can be considered as an end-directed (i. e., purposeful) process (Kondepudi et al., 2015).

The enactive approach (Stewart et al., 2010; Thompson, 2007) takes these kinds of precarious, self-maintaining systems as a conceptual starting point for defining agents (Barandiaran et al., 2009) and related phenomena, such as intrinsic normativity and teleology (Barandiaran and Egbert, 2013). But even if one does not subscribe to these approaches, the ability of these systems to adapt under seemingly arbitrary requirements (e. g., detecting the difference between these two frequencies of sound) makes these systems fascinating models for understanding the adaptability of biological organisms.

It is worth emphasizing the open-endedness of the adaptability of these systems. It would be difficult to argue that an acidic solution of dissolved ferrous sulfate has the inherent propensity to self-organize into a sound-discriminating ear, and yet by applying an electric potential across such a solution in a particular way, Pask & Beer were able to cause it to form into such a functional structure. This is rather remarkable. Does it hint at a not yet fully understood mechanism that might help us to understanding the remarkable open-ended adaptation demonstrated in nature? Cariani (1993, p. 20–21) suggests that by understanding the mechanism underlying Pask's result, we may come to understand how to create systems that can autonomously identify which features of the environment they respond to in a way that is more open-ended than that of conventional neural networks.

“[Conventional learning machines such as NN] improve on their (initial) designs by altering their decision functions contingent upon evaluation of past performance. But even with these machines, the designer must foresee the basic categories of percepts (i.e. primitive features) and actions which will be adequate to solve the problem at hand [...] For real world tasks, however, there is no such set of basic categories that is

defined beforehand, so that in addition to finding appropriate mappings there is also the problem of deciding what the basic categories will be. Essentially, contemporary trainable machines have the freedom to adapt within a set of percept and action categories, but they do not have the freedom to modify those categories. [...] Pask was specifically looking for a machine that would create its own “relevance criteria”, one which would find the observables that it needed to perform a given task. The device [would develop] sensors to choose, independent of the designer, those aspects of its external environment to which it would react. Not only would particular input-output combinations be chosen but the categories of input and of output would be selected by the device itself.”(Cariani, 1993, p. 21)

Some might argue that modern neural networks are capable of selecting their own categorisation schemes, but even if this is granted, they do not (yet) innovate a sensor that wasn't there before.

To summarize, the system studied by Pask and Beer (and the related RCTN structures) are worthy of further study as they (i) demonstrate unusual dynamics; (ii) are comparable to the precarious adaptation of individual biological organisms; and (iii) demonstrate an apparent open-ended ability to adapt. In the next section, we introduce a computational model of Pask and Beer's system. In the following section, we describe our efforts to use a reward-like variation of the applied voltage so as to steer the growth of the dendritic structures.

The reward function that our model optimises is considerably simpler than the task that Pask and Beer set for their system, but our model nevertheless demonstrates a plausible mechanism by which a physical system of this kind can respond to reinforcement signal at all. Our model builds to some extent on ideas presented in [cite Virgo and Harvey 2008], but the mechanism is much more physically realistic.

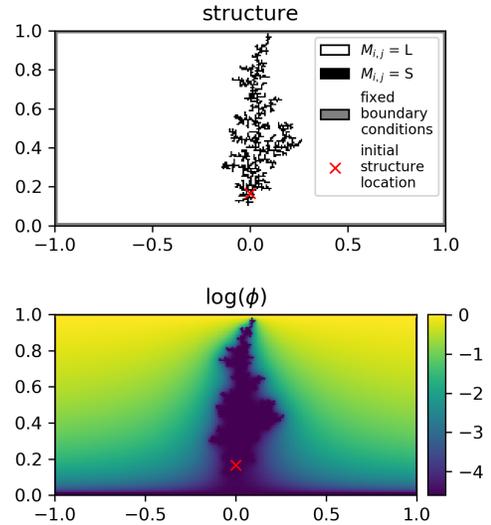
## Model

We now present a model of electrochemical deposition. Using finite difference methods, we use a rectangular  $256 \times 128$  lattice to simulate a two dimensional space 2 units wide by 1 unit tall. Each position on the lattice is considered to be either a negatively-charged highly conductive solid or an insulating liquid,  $M_{i,j} \in \{S, L\}$ . The electric potential,  $\phi$ , is calculated across this lattice by fixing the conductive solids (i. e., treating them as boundary conditions) and then solving the Dirichlet problem for the Laplace equation,

$$\nabla^2 \phi = 0 \quad (1)$$

by numerically integrating (using the forward-time centered space method —see e. g., Recktenwald, 2004) the heat equation,

$$\frac{\partial \phi}{\partial t} = \nabla^2 \phi \quad (2)$$



**Figure 2: Example final state of a simulation trial, showing the grown structure (top) and its surrounding electric potential field (bottom).**

until the system has come (close) to equilibrium i. e., until  $\forall \phi : |\phi_n - \phi_{n-1}| < 10^{-3}$ , where the subscript  $n$  is the current iteration index. When approximating the initial equilibrium for any given run of the simulation, we increase accuracy by reducing the tolerance by an order of magnitude, i. e.,  $\forall \phi : |\phi_n - \phi_{n-1}| < 10^{-4}$ .

The boundary conditions for the Dirichlet problem vary between experiments. We describe this variation in detail below, but in every case, there is a negatively charged conducting solid with a fixed relative potential of 0. We call this ‘the structure’ and it grows and decays via simulated electrochemical deposition and dissolution as described below. In addition, each experiment also always includes a positive boundary condition, that corresponds to an electrode with a fixed relative positive charge.

Each iteration of our simulation begins by approximating the electric-potential field equilibrium as described above. We then determine how the conducting structure will grow or decay. To do so, we identify  $I$ , a set of ‘interface cells’: liquid locations with one or more solid locations in their von Neumann neighbourhood.

$$I = \{(i, j) : M_{i,j} = L \wedge \exists S \in \{M_{i-1,j}, M_{i+1,j}, M_{i,j-1}, M_{i,j+1}\}\}. \quad (3)$$

For each interface cell, we calculate the probability that it will become part of the conducting structure. The probability of these ‘constructive’ changes are proportional to  $\phi$  of the interface cell (as  $\phi$  is the potential relative to the structure, which is proportional to electron flow at the interface).

To calculate this probability, use the following equation,

$$P_c = \left\{ \frac{1}{Z} (\phi_{i,j} + \zeta) : (i,j) \in I \right\}, \quad (4)$$

where  $\frac{1}{Z}$  is a normalization factor selected such that the sum of all of the probabilities is 1 and  $\zeta$  is a parameter that scales the relative influence of the voltage compared to entirely random process—as  $\zeta$  approaches infinity, the probabilities become equal across the interface cells. The results of varying  $\zeta$  can be seen in Figure 3. Essentially, as  $\zeta$  is increased, the structure that grows loses its filamentous structure and becomes less sensitive to  $\phi$ -gradients, i. e., stops growing toward high values of  $\phi$ .

Destructive events, where one of the solid neighbours of the interface cell becomes part of the liquid insulating material are also possible. The probabilities of the destructive events are a function of  $a(i,j)$ , defined as the mean age (time since creation) of the solid cells in the Moore neighbourhood of the interface cell. The assumption underlying this distribution is that over time, the existing structure become smoother and thus less likely to dissolve. This age-based probabilities are calculated according to the following equation.

$$P_d = \left\{ \frac{1}{(1 + a(i,j))^2} : (i,j) \in I \right\}. \quad (5)$$

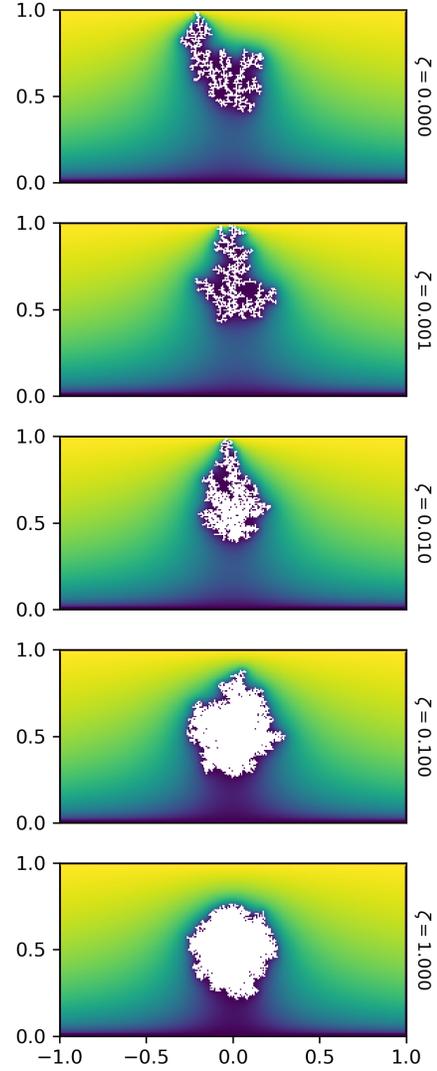
In Pask and Beer’s experiments voltage was varied so as to reward (i. e., stabilize / cause to grow) or punish (destabilize / dissolve) the structure. In our model we similarly have a reward parameter,  $r \in [0, 1]$  that biases the relative likelihood of constructive vs. destructive events. The next event thus selected from the following set:

$$P = \left\{ \frac{rp}{\sum P_c} : p \in P_c \right\} \cup \left\{ \frac{(1-r)p}{\sum P_d} : p \in P_d \right\}. \quad (6)$$

It is important to note that the reward function varies over time (as a function of system state) but not over space. As we shall see the structure tends to grow rather directly toward regions of high  $\phi$ , but it possible to counterdict this energy-seeking behaviour by selectively rewarding certain types of growth and punishing others by varying  $r$ .

Once the event is identified, the structure grows or decays as appropriate, the electric potential equilibrium is recalculated,  $I$  is updated, the probabilities for the next event are calculated etc., in a repeating iterative manner. Because of the probabilistic selection of events, each iteration corresponds to a different amount of time passage, specifically:  $\Delta t$  is taken from an exponential distribution with the rate parameter,  $\lambda$ , is the sum of the scaled but not normalized probabilities, i. e.,:  $\lambda = r \sum P_c + (r - 1) \sum P_d$ .

The stochastic degradation of the structure means that it is not uncommon for sections of the structure to become disconnected. When this happens, sections that are not contiguous with the initial starting point of the structure (which



**Figure 3: Increasing  $\zeta$  results in fewer dendrites and reduced electro taxis.** A seed structure at  $(0, 0.5)$  responds to a voltage gradient ( $\phi_{y=0} = 0, \phi_{y=1} = 1$ ) in different ways depending upon the relative influence of randomness and voltage as described by simulation parameter  $\zeta$ .

is interpretable as the negative electrode) are assumed to fall to a neutral potential and rapidly dissolve. This is simulated by removing any solid structure cells that are not connected to the initial starting position of the structure. This is not a physically realistic aspect of our simulation, but rather a simplification. In future work we may model disconnected conducting elements in a more realistic manner. Finally, we make it impossible for the first ‘seed’ cell of the structure to dissolve.

## Experiments & Results

We now evaluate different strategies for modulating the reward signal,  $r$ . Each experimental reward strategy is a function of  $x^*$ , the current mean horizontal position of the top-most part of the structure. To calculate this value, we identify the top-most row of  $M$  that contains structure  $i^* \equiv \max_i |M_{i,j} = S$ , and calculate the mean  $x$  of those positions within this row that contain structure,  $x^* \equiv \bar{J}$  where  $J \equiv \{ \frac{2j}{256} - 1 | M_{i^*,j} = S \}$ .

In each case the structure is seeded at the middle of the space, close to the bottom  $(x, y) = (0, \frac{1}{6})$ . A fixed voltage  $\phi = 1$  electrode is simulated as spanning the top edge of the area. To impose a gradient, the bottom edge is clamped to a value of  $\phi = 0$  and the left and right boundaries are also clamped as a linear gradient between the top and bottom boundary conditions. Formally:  $\phi_{y=1} = 1$ ;  $\phi_{y=0} = 0$  and  $\phi_{x=-1} = \phi_{x=1} = x$ . In all of the following experiments, we simulate the growth of these structures until either the structure touches the top electrode, or 25,000 iterations have passed. For each of the following experiments, we fix  $\zeta = 0$ .

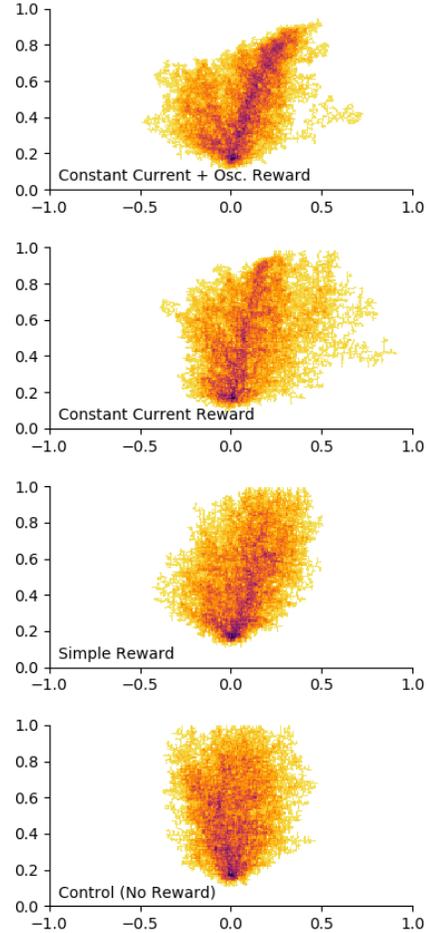
**Control Conditions.** We will first describe the control strategy where reward is fixed at  $r = 1$ . An example of the type of structure that grows in the control condition can be seen in Figure 2. The initial seed, located at  $(0, \frac{1}{6})$  grows rapidly and rather directly to the positive electrode. The bottom frame of Figure 4 shows the average density of 25 structures grown in these conditions. In every case, the trial ends when the structure has grown to the top of the simulation, and the horizontal location of the top of the structure is distributed approximately evenly around the centre of the arena (see the bottom row of Figure 6).

**Strategy 1: Simple Reward.** The first experimental reward strategy we consider is the modulation of  $r$  according to the following simple linear function of  $x^*$ . Our goal here is to encourage the structure to grow to the right.

$$r = mx^* + b \quad (7)$$

It is not self-evident which values of parameters  $m$  and  $b$  will maximise our influence of the structure. Figure 5 shows the conducted a systematic survey to investigate the influence of these parameters. Of the values tested, the parameters that maximised the mean rightward growth of the structures were  $(m, b) = 5.06, 0.5$  and it is these parameters that were used to generate the ‘Simple Reward’ portions of Figures 4 and 6.

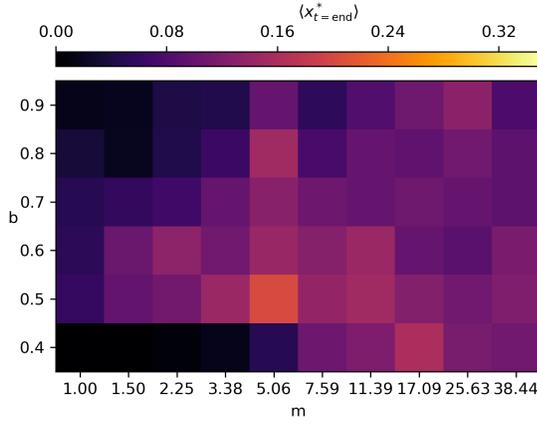
This simple strategy succeeds at influencing the growth of the structure. Once the structure has grown a little bit to the right of its initial location,  $r$  increases, and provided the structure does not grow back to the left,  $r$  will remain high enough for the structure to continue to grow. In other words, after an initial growth to the right, further rightward growth



**Figure 4: Superposition of the final states of 25 trials for the best found parameterization of each indicated strategy.** Darker areas indicate locations where structure existed at the end of a greater number of trials.

is unnecessary for the system to grow, and so it grows, attracted by the higher values of  $\phi$  close to the top of the arena. Is it possible to do better?

**Strategy 2: Constant Current.** Pask generally refers to current rather than voltage when describing this experiment, and so there has been some speculation that they used a constant current device that regulates voltage so that the total current flowing between the two electrodes is constant (or kept below some maximum)—see e.g. the description in (Bird and Di Paolo, 2008, p. 201). As conductive structure grows between the electrodes, resistance decreases. If the applied voltage were fixed, the current flowing through the system would thus increase as the resistance dropped. The constant current regulator is a simple feedback control device that regulates the applied voltage so that current is con-



**Figure 5: Systematic survey of simple-reward parameters.** Values indicate the mean  $x^*$  found for the indicated parameter values, with brighter values corresponding to greater influence of growth.

stant. Essentially if the current is greater than a target value, the applied voltage is decreased and vice-versa according to negative feedback-like relationship similar to  $\frac{dV}{dt} = \theta - A$ , where  $V$  is the voltage,  $\theta$  is the target current (selected by Pask) and  $A$  is the current.

For our second experimental reward strategy we simulate a constant-current feedback mechanism so that current is maintained at a target value ( $\theta$ ),

$$r = \frac{1}{2} + \theta - A, \quad (8)$$

where  $A$  is the total current between the positive electrode and the structure, which is calculated by summing the  $\phi$  values at the interface,

$$A = \sum_{(i,j) \in I} \phi_{i,j}. \quad (9)$$

The target current ( $\theta$ ) is then dynamically modulated according to how far right the structure has grown (again approximated by  $x^*$ ), thus:

$$\theta = \frac{1}{2} + m \left( x^* - \frac{1}{2} \right). \quad (10)$$

Once again it is not clear how to select a value for parameter  $m$  in this strategy. We experimented with  $m \in \{10, 20, 30, 40, 50, 60\}$ . The most effective value at maximising the mean  $x^*$  was  $m = 40$  maximised, and so we used that parameter value to generate the data plotted for this strategy in Figures 4 and 6. In these figures we can see that the constant current strategy produces a distribution of  $x^*$  with a mean that is greater than the simple strategy, but note that this is not a statistically significant increase (see Table 1).

	Simple	Const. Curr.	CC + Osc.
Control	$p < \mathbf{0.001}$	$p < \mathbf{0.001}$	$p < \mathbf{0.001}$
Simple		$p < 0.141$	$p < \mathbf{0.006}$
Const. Curr.			$p < 0.609$

**Table 1: Tukey’s test.** This table indicates chance that variation between means of the data plotted in Figure 6 is due to chance. Bold entries are considered statistically significant.

**Strategy 3. Constant Current with Exploratory Oscillations.** While observing the simulations of the constant current strategy, there was often a feeling of wishing that the structure would ‘experiment’ more—i. e., try out different random configurations and keep those that increase the reward signal. To encourage this kind of exploration, we added an externally driven oscillation to the reward signal to produce our final strategy.

The reward function is the same as in the Constant Current strategy, except that we update Equation 8 to include a sinusoidal function of the current iteration of the simulation,  $\tau$ . It would be more appropriate to have this be a function of time rather than iteration, and this will be an improvement that we make in future work.

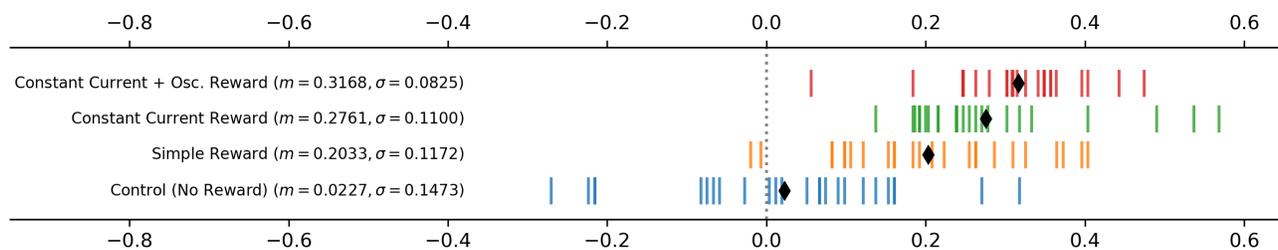
$$r = \frac{1}{2} + \theta - A + n \cos(2\pi\tau/p), \quad (11)$$

Once again, the control strategy includes free parameters, and we used a systematic survey to search for those that are more effective. Figure 7 show the results of this survey. There is no clear trend among these parameters, but they all perform well compared to the previous reward strategies. We selected the best performing parameters  $(n, p) = (0.1, 100)$  to generate the data plotted in Figures 4 and 6. The distribution of  $x^*$  generated by this strategy now significantly outperforms the simple reward mechanism ( $p < 0.001$ —see Table 1).

## Observations & Discussion

We have presented a new model for exploring the electrochemical deposition system investigated by Pask & Beer in the 1950s. The model has helped us to understand how by the selective rewarding of particular patterns of growth, it is possible to influence or ‘steer’ the dissipative structures that grow in these conditions. This is a potentially significant result, because it suggests a novel, and simple, mechanism through which physical systems can respond to reinforcement signals, potentially producing complex, organised structures as a consequence.

Our ability to guide the form of these structures in our model is not absolute and it is interesting to consider the source of any limitations and thus how they might be overcome. One limitation may come from the energetic gradients inherent in our simulation whereby the conducting structure naturally grows up  $\phi$  gradients. Each reward strategy rewarded growth orthogonal to the  $\phi$  gradients, but the



**Figure 6: Comparison of best parameterization for each evaluated strategy.** Each row indicates the  $x$  position of the top-most structure element for 25 trials of indicated reward strategy.

structures all also (unsurprisingly) responded primarily to the  $\phi$  gradient by growing upwards. Decreasing the influence of the  $\phi$  gradient might be expected to improve the steer-ability. One way to do this would be to increase  $\zeta$ . In the physical experiment, this would correspond with decreasing the relative voltage between the electrodes. But decreasing the voltage excessively would mean that no deposition would occur. A constant current mechanism might allow the voltage to remain high, while decreasing the ‘attractive force’ of the positive electrode. To speculate: the constant current mechanism partially neutralizes the attractiveness of the  $\phi$  gradient, as growth toward the positive electrode reduces resistance, which increases the current, which would cause the constant current mechanism to decrease the applied voltage. If properly tuned, such a mechanism might mean that the growth of the structure responds *only* to the reward function (and not also to the  $\phi$  gradient as is currently the case). In this way, a well-designed reward mechanism would ‘flatten’ the landscape of possible structures, facilitating the growth of those structures that maximise the reward signal.

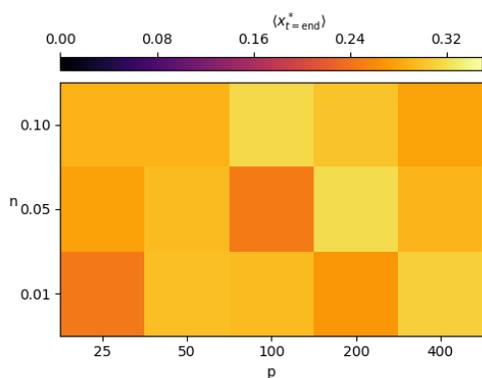
It is also interesting to consider this ‘flattening’ of the search space in the context of genetic evolution, where the search space of nucleotide sequences is essentially flat (i. e.,

there is little inherent cost for choosing a adenine or a guanine), facilitating the ability of evolution to search the space of polypeptide sequences unabated.

**Comparison to evolution.** In the introduction, we compared random but selected growth of the structure to the selection of random mutations in Darwinian evolution. Metaphors like this are useful both for identifying similarities between systems, and for highlighting differences. One such difference that we noticed in the simulation is that when a new branch begins to grow, it tends to grow in that same direction for some time. This inertia-like effect may be due to the tendency of new structures to grow into areas where they are more exposed, and thus subject to higher voltage and thus more likely to grow further—a kind of autocatalytic growth. It may be that increasing the reward signal during one of these may further accelerate this tendency allowing for a more instructive or directive kind of reward mechanism along the lines of “do more of that” rather than the post-hoc reward, “what you just did was good, keep it”. Darwinian evolution has no explicit inertia mechanism such as that just described, but it is interesting to reverse the metaphor and to consider that there are occasions when a new mutation opens up a set of possible environmental interaction which encourages further mutations.

**Comparison to metabolism and biological individuality.**

It is also interesting to consider one of these structures as a model of a biological individual performing a metabolism-based behaviour. The dendritic structure is a dissipative structure that relies upon the dissipation of energy (the flow of electricity) to persist. It reconfigures itself to amplify or stabilize this flow of energy, and when conditions are right, this adaptation can respond, not just to physical energy gradients (control case), but to more complex requirements (shown here in simulation and in Pask and Beer’s original experiments). A number of other dissipative structures similarly act so as to satisfy their own needs —see e. g. RCTN discussed in introduction and motile oil-droplets (Hanczyc, 2011). These physical systems, like some bacteria (Egbert et al., 2010) are responding essentially to their own rates of self-construction in what is referred to as metabolism-based



**Figure 7: Systematic survey of simple-reward parameters.** Values indicate the mean  $x^*$  found for the indicated parameter values, with brighter values corresponding to greater influence of growth.

behaviour, which can facilitate adaptation and evolution in a number of ways (Egbert et al., 2012; Egbert and Pérez-Mercader, 2016).

**Origins of life.** Pask’s electrochemical experiments seem to demonstrate the emergence of functional components (for example, the vibrating filaments in the ‘ear’), without them needing to be ‘designed in’ by a human engineer. One can find examples of this in other places, such as the emergence of new traits in evolution, or the emergence neurons that perform specific filtering operations when training a deep neural network, for example. But Pask’s case is remarkable because it occurs in such a simple physical system, purely as the result of the physical processes of fluctuating growth and decay of filaments, in response to a reward function.

Although there are some parallels with evolution, this simple physical mechanism differs in that there is no need for large, complex polymer molecules to be produced. If evolution is not the only mechanism by which complex functional structures can arise in the natural world, it becomes possible that evolution as we know it is the *result* of a dynamical process, and not just its starting point. The emergence of Paskian growth seems much easier than the emergence of complex biomolecules, and perhaps mechanisms resembling it played a role in steering the abiotic world on its way to the emergence of biology.

**Self-organizing steerable self-organizing systems** In our model there are a number of parameters that would influence the extent to which the self-organization is controllable or ‘steerable’. That Pask and Beer were capable of finding conditions suitable for steering the self-organization of an ‘ear’ is remarkable. Instead of trying to identify effective parameter regimes directly, it may be more effective to identify negative-feedback-like mechanisms that automatically regulate parameter regimes so as to produce steerable self-organization. One example of such a mechanism may be the constant-current mechanism as described above.

**Future work** We have proposed a mechanism through which a simple electrochemical system could plausibly respond to a reward signal. However, the task we set our system, of growing in a particular direction, was very simple in comparison to the task of “growing an ear” that was achieved by Pask and Beer. It will be important in future work to show that this kind of system can solve more difficult tasks. It will be equally important to build a better theoretical understanding, in order to understand whether other kinds of physical system can exhibit similar reinforcement learning behaviour.

## Acknowledgements

Preparation of this manuscript was generously supported by a Royal Society Te Apārangi, Marsden Fund Te Pūtea Rangahau A Marsden grant awarded to Matthew Egbert in 2017 (Project ID: 17-UOA-196).

## References

- Barandiaran, X. E., Di Paolo, E., and Rohde, M. (2009). Defining Agency: Individuality, Normativity, Asymmetry, and Spatio-temporality in Action. *Adaptive Behavior*, 17(5):367–386.
- Barandiaran, X. E. and Egbert, M. D. (2013). Norm-Establishing and Norm-Following in Autonomous Agency. *Artificial Life*, 20(1):5–28.
- Bird, J. and Di Paolo, E. (2008). Gordon Pask and His Maverick Machines. In Husbands, P., Holland, O., and Wheeler, M., editors, *The Mechanical Mind in History*, pages 185–211. The MIT Press.
- Boden, M. A. (2010). *Creativity and Art: Three Roads to Surprise*. Oxford University Press. Google-Books-ID: MS0UDAAAQBAJ.
- Cariani, P. (1993). To evolve an ear. Epistemological implications of Gordon Pask’s electrochemical devices. *Systems Research*, 10(3):19–33.
- Egbert, M. D., Barandiaran, X. E., and Di Paolo, E. A. (2010). A Minimal Model of Metabolism-Based Chemotaxis. *PLoS Computational Biology*, 6(12):e1001004.
- Egbert, M. D., Barandiaran, X. E., and Di Paolo, E. A. (2012). Behavioural Metabolism: The adaptive and evolutionary potential of metabolism-based chemotaxis. *Artificial Life*, 18(1).
- Egbert, M. D. and Pérez-Mercader, J. (2016). Adapting to Adaptations: Behavioural Strategies that are Robust to Mutations and Other Organisational-Transformations. *Scientific Reports*, 6:18963.
- Hanczyc, M. M. (2011). Metabolism and Motility in Prebiotic Structures. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1580):2885–2893.
- Huynen, M. A. (1996). Exploring phenotype space through neutral evolution. *Journal of Molecular Evolution*, 43(3):165–169.
- Jun, J. K. and Hübner, A. H. (2005). Formation and structure of ramified charge transportation networks in an electromechanical system. *Proceedings of the National Academy of Sciences of the United States of America*, 102(3):536–540.
- Kondepudi, D., Kay, B., and Dixon, J. (2015). End-directed evolution and the emergence of energy-seeking behavior in a complex system. *Physical Review E*, 91(5):050902.
- Nicolis, G. and Prigogine, I. (1977). *Self-organization in nonequilibrium systems: from dissipative structures to order through fluctuations*. Wiley, New York, NY.
- Pask, G. (1958). Physical analogues to the growth of a concept. pages 765–794, National Physical Laboratory, London. H.M.S.O.
- Pask, G. (1959). The Natural History of Networks. In Marshall C. Yovits and Scott Cameron, editors, *Self-Organizing Systems: Proceedings of an Interdisciplinary Conference*, volume 2, pages 232–263. Pergamon Press.
- Recktenwald, G. W. (2004). Finite-difference approximations to the heat equation. *Mechanical Engineering*, 10:1–27.
- Stewart, J., Gapenne, O., and Di Paolo, E. A. (2010). *Enaction: Toward a New Paradigm for Cognitive Science*. MIT Press.
- Thompson, E. (2007). *Mind in Life*. Harvard University Press, London, England.