

Ionic Electroactive Polymer Hydrogels Exhibit Emergent Learning Behaviour When Embodied in a Simulated Game-World

Vincent Strong^a, William Holderbaum^a, and Yoshikatsu Hayashi^{a,1}

^aDepartment of Biomedical Sciences/Engineering, School of Biological Sciences, University of Reading, Reading, UK

This manuscript was compiled on February 18, 2023

1 The goal of artificial neural networks is to utilise the structure of the
2 biological brain to enhance computational algorithms, however these
3 implementations still cannot achieve the kind of emergent learning be-
4 haviour found in Biological Neural Networks (BNNs). Reservoir com-
5 puting is an ever-growing field that is expanding what was thought
6 capable in computation through the use of unexpected computational
7 mediums. Mediums from slim mold to buckets of water and even
8 biological neurons, have been used to perform computation over and
9 beyond that of traditional digital computers. These organic computers,
10 that integrate organic neurons with computer hardware, are capable
11 of the kind of emergent learning found in BNNs. By applying the theo-
12 ries of learning in BNNs, such as active inference via the Free Energy
13 Principle (FEP), to learning within a different system whose behaviour
14 is also governed by FEP, can similar emergent learning be achieved
15 with a more accessible medium? In this paper ionic Electro-Active
16 Polymer (EAP) hydrogels were embedded in a simulated game-world,
17 based on the game Pong. The hydrogel was interfaced with a com-
18 puter system via a custom built Multi-Electrode Array (MEA). Through
19 analysis of the behaviour of the EAP hydrogel within the simulated
20 game environment, improved performance was observed over the
21 course of the game. Interesting emergent behaviour was also ob-
22 served in the hydrogel's reaction to environmental information. These
23 observations enforced the theory of FEP, its importance to learning
24 within BNNs and its ability to allow learning in other mediums, when
25 the current interfacing system is implemented.

EAPs | Emergent Learning | Reservoir Computing

1 **E**mergent computing is a field inspired by the computation
2 that takes place throughout nature. Emergent Computing
3 is described as highly complex processes arising from the co-
4 operation of many simple processes (1), or when behaviour of
5 a system does not depend on its individual parts but on their
6 relationships to one another (2). Artificial neural networks
7 were developed based on the emergent computing behaviour of
8 biological brains, both are large complex systems composed of
9 simple machines whose combined interactions lead to complex
10 computation or thought. The expansion of these theories led
11 to reservoir computing and the application of physical non-
12 linear systems, called reservoirs, as part of the computation
13 process. By using reservoirs found in nature, reservoir comput-
14 ing techniques allowed the emergent computation of nature to
15 be utilized in practical applications. For example with water
16 ripples for image analysis (3), mycelium mold as maze solvers
17 (4), and chemical reactors as logic gates (5, 6). These systems
18 solve calculation tasks and often provide unexpected solutions,
19 however they do not exhibit the emergent learning behaviour
20 observed with biological brains. Without learning behaviour a
21 reservoir systems capability to improve is limited, continuous

tasks that require behaviour based on task history are too
22 complex to yield any useful result. With the correct combi-
23 nation of task and computational medium, can the emergent
24 learning behaviour similar to that observed in biological neural
25 networks (BNNs) be achieved? Thus, allowing more complex
26 and nuanced tasks to be pursued through reservoir computing
27 techniques.

To harness the computational power of BNNs, for calcula-
28 tion problems, the field of Artificial Neural Networks (ANNs)
29 was developed by applying the learning capabilities of bio-
30 logical structures. ANNs have made great strides in solving
31 problems via machine learning that were previously thought to
32 be impossible to solve via computers. However, these solutions
33 are always limited by the hardware in which they are imple-
34 mented (7, 8). As approximations of BNNs, ANNs are not
35 capable of the learning found in their biological inspiration.
36 To better understanding how the interactions of biological
37 neurons allow for such complex useful behaviour, with the
38 goal of improving artificial neural network implementations,
39 theories of morphological computing and embodied cognition
40 have been employed (9, 10). Morphological computing theo-
41 ries explore the integration of computing behaviour with the
42 physical 'hardware' that produces it and proposes that the
43 computations a system performs are encoded within its struc-
44 ture (11). Similarly embodied cognition theorises that a body's
45 interactions with an environment constitutes or contributes
46 to the body's behaviour (12).
47

Significance Statement

Reservoir computing is an approach derived from recurrent neural networks receiving more and more interest due to its use of non-linear physical mediums, even incorporating BNNs to solve complex problems. Here we present a method for using a medium significantly more accessible than BNNs as a reservoir computing medium, EAP hydrogels. Utilizing theories in neurology and building on the behavioural similarities between ionic hydrogels and biological neurons, presented through the universal principles of free energy. We demonstrate this method by applying the EAP hydrogel reservoir to a task that has been approached by biological neuron based 'wet-ware' computers, showing how morphological computing techniques can exhibit the same extremely useful emergent learning behaviour as found in biological neurons when properly exploited.

V.S. and Y.H. designed research concepts and framework; V.S. performed experiments and developed computational platform; V.S. analysed data; V.S., Y.H. and W.H discussed the results and wrote the paper.

There are no competing interests associated with the contributing authors

¹To whom correspondence should be addressed. E-mail: y.hayashi@reading.ac.uk

48 to cognition, meaning that learning behaviour is a result of
49 a systems method of receiving information and thus a result
50 of its physical structure (12). Both these theories contribute
51 to the idea that to achieve equivalent learning behaviour of
52 BNNs in ANNs the hardware would need to be capable of
53 the kind complex interactions present in BNNs. This concept
54 of implementing complex interactions not possible in digital
55 hardware lead to the development of reservoir computing.

56 Reservoir computing is inspired by Recurrent Neural Net-
57 work (RNN) frameworks, where the dynamics of a fixed non-
58 linear system, called a reservoir, is used as part of a neural
59 network to map input and output signals to higher dimensional
60 space (13). This reservoir is typically a physical system that
61 exhibits complex behaviour that can be used to encode data
62 and embody computation not possible within the typical ANN
63 structure. A neural network utilising a reservoir consists of
64 three main layers, the excitation layer, the reservoir, and the
65 readout layer. The excitation layer converts the input to the
66 network into a form that can be used to stimulate or excite
67 the reservoir. The readout layer interprets the responses of
68 the reservoir into usable data, this layer also contains weights
69 allowing tuning of the interpretation to optimize the perfor-
70 mance of the reservoir neural network. The reservoir can be
71 any kind of medium that can encode temporal problems into
72 higher dimensions to generate recurrent connections between
73 data (14).

74 One of the most promising applications of reservoir com-
75 puting to achieve the capabilities of biological learning, is in
76 using biological neurons themselves (15). These applications
77 combine grown neuron cultures with electronic interfaces to
78 produce a computational device (16) referred to as organic
79 computing or 'wetware' computers. Biological neurons form
80 connections and alter weights based on feedback within the
81 network, connections that fire often are strengthened and
82 useful behaviour emerges. By stimulating BNNs with task in-
83 formation encoded into spikes the neurons will form pathways
84 in response to said task. Neural plasticity is the ability of the
85 nervous system to change its activity in response to intrinsic
86 or extrinsic stimuli by reorganizing its structure, functions, or
87 connections (17). The plasticity of a network can be altered
88 through the application of patterned or random stimulation
89 (18). By controlling this plasticity, the rate at which the net-
90 work reorganises itself can be altered, in this way a feedback
91 loop can be implemented where the network is 'punished' by
92 increasing plasticity when it performs the task badly or 're-
93 warded' by reducing plasticity when it performs well (18, 19).
94 There have been a few implementations of this technology,
95 such as in simulated flight control (20) or the arcade game
96 'Pong' (21).

97 In the paper "In vitro neurons learn and exhibit sentience
98 when embodied in a simulated game-world" (21) a grown
99 culture of neurons learn to play the game Pong in a virtual
100 environment. The game Pong consists of a ball and paddle,
101 the ball bounces around a frictionless 2D environment with
102 the goal of hitting the ball with the paddle as many times
103 as possible before it hits the wall behind the paddle, how-
104 ever the paddle can only move vertically. The game world
105 is encoded into stimulation by representing the location of
106 the ball as positional stimulation on the neural culture. An
107 area is then designated as a sensing region which is monitored
108 for responses, used as motor commands, to drive the paddle

109 position. These combined with a closed loop that manages
110 neural plasticity allows the BNN (referred to as DishBrain) to
111 improve in ability while playing the Pong game. The method
112 of controlled plasticity, that allows the BNN to learn, is derived
113 from theories of how intelligent behaviour arises in biological
114 brains via the free energy principle (FEP) (22–24).

115 FEP suggests that any self-organising system separate from
116 its environment seeks to minimise its variational free energy
117 (22, 25) by placing a bound on its long-term entropy (26). For
118 a given state there are a number of possible events, an entity
119 with agency inhabiting a state assigns probabilities to those
120 events. If the next event is known the probabilities distribution
121 is peaked, else the probabilities are equal and entropy of the
122 event distribution is maximised. FEP interprets learning
123 as an entity applying a bound to this maximum entropy or
124 surprise, limiting the free energy by entering states that alter
125 sensed information (24, 27). This creates a feedback loop
126 where internalised environmental information is continually
127 improved by actions that alter the environment, and can be
128 conceptualised by ordered interaction between local states
129 of the entity. The development of order in the interactions
130 of local states is known as self-organisation, where feedback
131 corrects deviations from an ordered configuration (28, 29).

132 Biological brains are a form of self-organising system, synaptic
133 pathways altering structure to allow for rules in state inter-
134 actions (30, 31). In this way the biological brain learns through
135 active inference, using an internal generative sudo model to
136 predict inputs that represent the external world (22, 25, 32).
137 The generation and continual updating of an internal model
138 to match external events is also explored through theories of
139 Bayesian inference, a common structure used to apply learning
140 in statistics (33) and interpret learning in biology (34). Put
141 simply, as sensed inputs, different from what are expected
142 given the internal model, are received the free energy increases
143 and the system seeks to reduce this free energy by the most
144 direct means possible. The reduction in free energy then
145 presents as the self-organisation of neural connections (35, 36).
146 Since the inputs influenced this change in structure the new
147 structure is a representation of the inputs, and by inference a
148 representation (or model) of the environment that provided
149 those inputs. This restructuring can present behaviourally
150 as either: action to reduce difference between the internal
151 model and sensed environment by making the environment
152 match the internal model, or by altering the internal model
153 to better reflect the environment. Under this theory, BNNs
154 hold 'beliefs' about the state of the environment, and learning
155 behaviour emerges by minimising internal free energy through
156 either, updating these beliefs, or take action to change the
157 environment to match these beliefs (25, 37). The aforemen-
158 tioned paper (21) showed these theories of learning to hold
159 merit, as their framework for applying the BNN to the Pong
160 game effectively allowed the BNN to perform and improve
161 in the game world (21). Although BNNs were shown to be
162 useful as a computation medium capable of learning in ap-
163 plication to complex tasks, BNNs are by definition biological
164 and because of this are subject to the difficulties involved in
165 working with organic materials. The foremost issue lies in the
166 fact that BNNs require organic neurons, these are difficult to
167 grow, maintain, and have a short shelf-life once synthesised
168 (38). However, FEP is the underlying mechanism leading to
169 emergent learning and BNNs are not the only system that

170 preset behaviour based on FEP.

171 Ionic Electro-Active Polymer (EAP) hydrogels are a form
172 of active matter material (39) that has shown promise in the
173 field of soft robotics as an actuator (40). Dynamics present in
174 soft bodies have also shown some potential as computational
175 resources (41) within soft robotics, presenting an interesting
176 potential for ionic EAPs as a computational resources given an
177 appropriate application and framework. As an EAP this ma-
178 terial changes shape in response to electric fields. As an ionic
179 EAP, the volume changes are induced through ion migration,
180 followed by osmotic pressure driven water flow. Hydrogen
181 ions act as active agents within the hydrogel, interacting with
182 the polymer networks. The dynamics of these active particles
183 have shown memory like behaviour (42), and the underlying
184 mechanisms can be explained through FEP. When an electric
185 field is applied this causes an increase in free energy, the ions
186 then seek to minimise this free energy by migrating to align
187 with the electric field (43). EAP hydrogels are considerably
188 easier to synthesise than biological neurons, requiring less
189 maintenance with much longer shelf life. Due to the simpler
190 method of synthesis, EAP hydrogels also allow for consider-
191 able versatility in physical form, allowing different shapes and
192 volumes to be produced with minimal additional complexity
193 to procedure. This accessibility when compared to biolog-
194 ical neurons makes EAP hydrogels an attractive alternative
195 in reservoir applications while taking advantage of the same
196 underlying mechanics. So, if the electric field provided to the
197 EAP hydrogel was representative of an environment would
198 the distribution of ions represent an internal model of the
199 environment, much like the structure of BNNs within the
200 brain? Furthermore, if ionic EAP hydrogels and BNNs exhibit
201 behaviour based of the same underlying principles of FEP,
202 then is it possible that EAP hydrogels can exhibit the same
203 learning capabilities if applied to an established BNN task
204 such as Pong?

205 In summary, there are many computation techniques that
206 are rapidly developing to achieve the kind of emergent learning
207 behaviour found in BNNs. Theories in embodied computation
208 have led to developments in ANNs, allowing physical medi-
209 ums with complex behaviours to be used as computational
210 resources (13) to overcome the digital hardware limitations.
211 Through reservoir computing BNNs were able to be integrated
212 into computational tasks by taking advantage of the under-
213 lying driving mechanics of FEP (21). Although this route
214 allows for the emergent learning of BNNs to be used, biolog-
215 ical neurons themselves are difficult to work with, requiring
216 specialist equipment to synthesise and having a short shelf-life
217 once produced among other issues (38). However, BNNs are
218 not the only medium whose behaviour is a result of FEP. Ionic
219 EAP hydrogels are also governed by FEP, ions migrating under
220 the influence of electric fields to reduce free energy of the
221 system (43). Ionic EAPs are significantly easier to construct
222 and maintain than biological neurons. If similar emergent
223 learning can be achieved with this more approachable medium,
224 the potential for use in computation would be far more impact-
225 ful and allow for more rapid development without specialist
226 equipment.

227 This study aims to show how EAP hydrogels can exhibit
228 emergent learning behaviour similar to that found in BNNs
229 when embodied in a simulated game-world. The game Pong is
230 used as the task as it has already been used to the application

of grown biological neurons as wet-wire computers (21). First,
231 by measuring ion concentrations through conductivity of the
232 EAP hydrogel after periods of stimulation, we studied the
233 memory mechanics of the ionic EAP hydrogel. Secondly, the
234 ionic EAP hydrogel is embodied in the simulated game world
235 of Pong through the use of a custom Multi-Electrode Array
236 (MEA). The game environment is encoded into stimulations
237 provided to the hydrogel and recorded ion concentrations are
238 used as motor commands within the game world. Multiple
239 game ‘runs’ are conducted with this setup and the performance
240 of the hydrogel over time within the simulated environment is
241 recorded and analysed.

Nomenclature

\bar{c}	Mean value of c , $\bar{c} \in \mathbb{R}^+$	244
ϵ	Absolute permittivity of material, $\epsilon \in \mathbb{R}^{>0}$	245
μ_U	Mean of U , $\mu_U \in \mathbb{R}^{>0}$	246
Ω	Number of microstates, $\Omega \in \mathbb{Z}^+$	247
Φ	Normal distribution	248
σ^2	Variance, $\sigma^2 \in \mathbb{R}^{>0}$	249
σ_U	Standard deviation of U , $\sigma_U \in \mathbb{R}^{>0}$	250
A	Surface area of electrode, $A \in \mathbb{R}^+$	251
c_i	Number of occurrences of the i^{th} element of O , $c_i \in \mathbb{Z}^+$	252
C_{cont}	Continuity adjustment, $C_{continuity} = -0.5$	253
C_{ties}	Adjustment due to tied ranks, $C_{ties} \in \mathbb{R}^{>0}$	254
D	Diffusion constant, $D \in \mathbb{R}^+$	255
d	Distance between opposing electrode plates, $d \in \mathbb{R}^+$	256
e	Electron charge, $e = 1.602 \times 10^{-19} C$	257
F	Faraday constant, $F = 9.649 \times 10^4 C \cdot mol^{-1}$	258
K	Set of unique ranks with ties, $K \in \mathbb{R}$	259
k	Rank, $k \in K$	260
k_B	Boltzmann constant, $k_B = 1.381 \times 10^{-23} m^2 \cdot kg \cdot s^{-2}$.	261
	K^{-1}	262
k_e	Coulomb constant, $k_e = 8.988 \times 10^9 N \cdot m^2 \cdot C^{-2}$	263
m_1	Concentration at ion, $m_1 \in \mathbb{R}^{>0}$	264
m_2	Concentration between electrodes, $m_2 \in \mathbb{R}^{>0}$	265
n	Size of dataset, $n \in \mathbb{Z}^{>0}$	266
O	Set of elements in Q without repetitions, $O \in Q$	267
o	Number of samples in O , $o \in \mathbb{Z}^+$	268
p	P value, $p \in \mathbb{R}^{>0}$	269
Q	Sample used in distribution, $\{Q \in \mathbb{Z} 0 \leq Q_i \leq 1000\}$	270
q	Number of samples in Q , $q \in \mathbb{Z}^+$	271
R	Gas constant, $R = 8.315 J \cdot mol^{-1} \cdot K^{-1}$	272
r	Distance of ion from electrode pair, $r \in \mathbb{R}^+$	273
R_s	Lowest ranked sum value, $R_s \in \mathbb{R}^{>0}$	274
S	Entropy, $S \in \mathbb{R}^{>0}$	275
T	Absolute Temperature, $T \in \mathbb{R}$	276
t_k	Number of ties for the k^{th} rank, $t_k \in \mathbb{Z}^{>0}$	277
U	U statistic, $U \in \mathbb{R}^{>0}$	278
u	Free energy, $u \in \mathbb{R}^{>0}$	279
v	Voltage of electrodes, $v \in \mathbb{R}$	280
v_1	Ion Voltage potential at starting location, $v_1 \in \mathbb{R}$	281
v_2	Ion Voltage potential between electrodes, $v_2 \in \mathbb{R}$	282
x	Distance over which ion motion is measured, $x \in \mathbb{R}^+$	283
Z	Charge of ion, $Z \in \mathbb{Z}$	284
z	Z score, $z \in \mathbb{R}$	285

Memory Mechanics Through Ion Migration and Conductivity Measurement

For a medium to be used in computation it must exhibit be-
288 haviours representative of properties key to computational
289

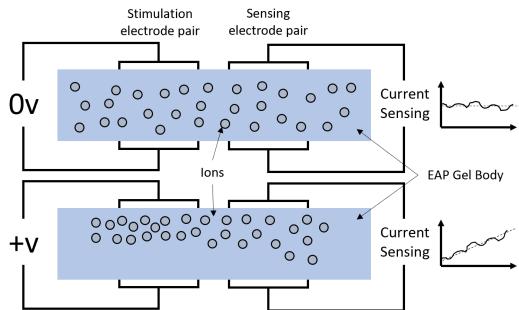


Fig. 1. Diagrams show how ion motion is influenced by external stimulation via an electric field and recorded via current draw. Under stimulation ions will gather between the electrode pair from elsewhere in the EAP gel. As the concentration of ions in this location increases the conductivity also increases, the increase in conductivity can be measured as an increase in current draw when a small voltage is applied in parallel close to this location via a secondary electrode pair.

system, the most fundamental of this is memory. Polyacrylamide hydrogel is an ionic EAP that displays mechanics that can be interpreted as a form of computational memory.

Polyacrylamide hydrogel is an active matter EAP (43), hydrogen ions act as active agents within the hydrogel influenced by the polymer network and each other. Stimulation by an electric field causes an increase in free energy within the hydrogel, the ions migrate (44) to minimise the free energy of the system (43). As the ions move, they drag water molecules causing changes in water distribution and localized deformation of the gel, driven by an equilibrium between osmotic pressure (45) and rubber elasticity (46) in the polymer network. Given a constant migration of ions into a location the localised rate of swelling gradually decreases, further decreasing free energy, this creates a hysteresis effect (47). This hysteresis effect, leads to complex dynamics in the gel's reactions to stimulation.

The hydrogen ions take relatively little time to migrate under stimulation by an electric field, but take considerably longer to diffuse to a homogeneous distribution under no stimulation. The difference in time scale allows previous stimulations to affect future stimulations, as the ion distributions persist between stimulations leading to a form of memory (42, 48).

This experiment is designed to demonstrate this memory mechanic, and detail the underlining methodology that is used in the subsequent experiment presented in this study. By recording the ion concentrations before and after consecutive stimulations the "remembered" state caused by the stimulation can be observed. Ion concentration can be measured through the conductivity of the gel. As ions collect in a location and ionic concentration increases, the local conductivity also increases (49). By applying a small voltage to said location, the current draw on the power supply is directly proportional to the conductivity and thus directly proportional to the ion concentration. By utilizing two adjacent pairs of electrodes stimulation can be applied while conductivity is simultaneously measured, this setup is illustrated in figure 1.

To better understand the behavior within the hydrogel EAP system, as shown in figure 1, the basic mechanisms at play can be described through Frick's first law of particle diffusion (50). Frick's first law can be used to describe diffusive flux J over a free energy gradient δu , shown in equation 1. There are three main components to the free energy of ions within the hydrogel system u , shown in equation 2; energy due to the

electric field imposed by the electrode pair u_e , energy imposed by the changing chemical concentration gradient within the hydrogel as ions migrate u_c , and energy due to the electrical potential gradient of ions within the hydrogel as ions migrate u_p . u_e and the electrochemical gradient $u_c + u_p$ oppose each other, as the electric field attracts ions to a location, the local concentration and potential increases causing an increase in free energy opposing that of the electric field.

$$J = \frac{D}{RT} \frac{\delta u}{\delta x} \quad [1]$$

Where:

$$\delta u = u_e + u_c + u_p \quad [2]$$

$$u_e = k_e \frac{v \epsilon A e}{dr} \quad [3]$$

$$u_c = RT \ln \left(\frac{m_2}{m_1} \right) \quad [4]$$

$$u_p = ZF(v_1 - v_2) \quad [5]$$

u_e can be represented via Coulomb's law and electric potential energy, shown in equation 3. Ion charge is equal to electron charge, the charge imposed by the electrode pair is derived from the capacitance of the electrodes surrounding the hydrogel. u_c can be derived from the Gibbs free energy representation of entropy as the ratio between concentration increases (51), shown in equation 4. u_p is represented via the Gibbs free energy derivation of cell potential as the combined charge of ions becomes inhomogeneous (51), shown in equation 5.

Due to Coulomb's law, as the ions approach the charged electrode pair the u increases resulting in a larger flux J , however $u_c + u_p$ also increases opposing the electric field. These two opposing forces change at very different rates, the difference in magnitude being a key factor of the hysteresis component. As the ions migrate under the influence of the electric field, eventually the u is minimised as the opposing factors balance. This interpretation is however a simplification, in actuality there are many additional free energy components that contribute to the complexity of the hydrogel's EAP behavior such as terms associated with the deformation the gel's polymer structure as it swells due to local ion and water concentrations. These additional terms also influence factors of other terms such as diffusion coefficients and permittivity of the hydrogel material, resulting in extremely complex behavior as the free energy is minimised.

Experimental Design. Using this setup, a series of stimulations were applied to a polyacrylamide hydrogel sample to demonstrate the memory mechanism. Polyacrylamide is a relatively simple EAP hydrogel to synthesise (52), allowing for it to be produced in large batches and moulded in many different ways for the required experiments. The gels were synthesised using the methodology described in the supplementary information section "Gel Synthesis Procedure" based on the procedure detailed in (43). The gel was left in a sodium chloride solution (0.08%) to allow the ion concentrations to stabilize and provide consistent starting water concentration within the hydrogel. The gel was then placed between two electrode pairs spaced with a 7mm gap between them. A constant sensing voltage of 2 V was applied to one electrode pair, using a current sensor

381 to record the response, while a series of 20 V stimulations
382 were applied to the second electrode pair, both voltages were
383 applied with the same polarity.

384 Temperature and water content within the hydrogel can
385 affect free energy equilibrium (43). Temperature alters the
386 elastic properties and rate of ionic motion (52). As the osmotic
387 pressure difference between the polymer networks and ionic
388 solution drives the swelling, changes in the hydrogel's water
389 content change the degree to which it can swell, as well as alter-
390 ing the mechanical properties of the material (53). Because
391 of this temperature and water content must be controlled. To
392 mitigate this as much as possible the water was taken from a
393 temperature controlled source (approximately 22C) to ensure
394 the solution was a consistent temperature and provide the
395 same starting temperature conditions. Furthermore, the gels
396 were stored in water tight containers and left in the ionic
397 solution for the same period of time to ensure the starting
398 water content of the hydrogels were consistent.

399 **Results.** The results of this demonstration can be seen in figure
400 2. The recorded current has a default value of approximately
401 0.8 mA as the EAP hydrogel is inherently conductive shown
402 in segment A. With application of an electric field there is
403 an immediate rise to 2 mA as the field is picked up by the
404 sensing electrode pair shown at the beginning of segment B.
405 The stimulation also causes an increase in free energy, as a
406 result the ions move toward the electrode pair to minimise free
407 energy. The current draw rises, as ions collect and conductivity
408 rises, to 2.4 mA shown in segment B. Once the stimulation is
409 removed the current drops back to default values as shown in
410 segment C. Once the stimulation is reapplied the current draw
411 immediately rises back to the point where segment B ended as
412 the ions still maintain their positions, as shown at the start of
413 segment D. This observed behaviour exhibits a memory like
414 mechanism where a value is "saved" to the EAP hydrogel by
415 the redistribution of ions within the gel in response to the
416 increase in free energy. This behaviour can then be seen again
417 at the beginning of segment F where it continues from where
418 segment D finished.

419 **Discussion.** From this demonstration it is evident that,
420 changes in the ion distribution and polymer network can
421 be recorded as current draw through the application of small
422 voltages, evident by the change in current over time as stim-
423 ulation is applied altering the ion distribution and polymer
424 network. The demonstration also highlights how the increase
425 in free energy through stimulation is resolved by the hydrogel
426 through ion migration, and the last state of the EAP hydrogel
427 "remembered" if stimulation is removed. With reapplication of
428 the stimulation the changing current continues from the last
429 state recorded during stimulation. From these observations
430 it can be established that a form of memory is present in
431 the EAP hydrogel, and the response due to the hydrogel's
432 memory can be measured in parallel with the application of
433 stimuli. To fully demonstrate this memory mechanic and its
434 value within computation a suitable activity is required, along
435 with a closed loop control structure and hardware to interface
436 with the hydrogel.

437 **EAP Hydrogel Embodied in a Simulated Game-world**

438 The behaviours of the EAP hydrogel can be characterized
439 through free energy minimisation (43). There are many sys-

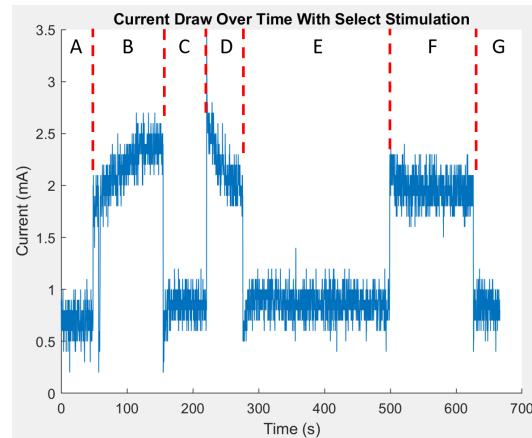


Fig. 2. This graph shows a sample of recorded current draw from 2 adjacent parallel electrodes as described in 1. The data has been segmented and labelled A through G for reference. The sensing voltage (2 V) was applied continually through the entire demonstration, stimulation (20 V) is applied in sections B, D, and F. The recorded current has a default value of approximately 0.8 mA as the EAP gel is inherently conductive shown in segment A. With application of an electric field there is an immediate rise to 2 mA as the field is picked up by the sensing electrode pair shown at the beginning of segment B.

tems whose mechanics can be represented through free energy
440 minimisation, regarding computation, one of the most sig-
441 nificant is that of learning within BNNs (22). Furthermore,
442 in the most basic sense, biological brains establish memory
443 through the arrangement of neurons (54) much like the EAP
444 hydrogel does through ion distributions. With the memory
445 mechanics present in the EAP hydrogel demonstrated, the
446 computational learning can be quantified through an activ-
447 ity that has also been used to assess biological neurons in
448 computing applications.

The paper "In Vitro Neurons Learn and Exhibit Sentience
450 When Embodied in a Simulated Game-World" (21) used the
451 game Pong as the quantifying activity when integrating a
452 BNN with computer architecture. The game provided a vir-
453 tual environment for the BNN to inhabit and learn from. The
454 method used to interface the biological neurons with the sim-
455 ulated environment fits neatly with the reservoir computing
456 framework (13). In the paper (21) the position of the ball
457 was presented to the neurons through localized stimulation,
458 aligning with the function of the excitation layer in a reservoir
459 computer, and the response of the neurons were converted
460 to motor commands to control the paddle, aligning with the
461 function of the readout layer. Given the commonalities in
462 the underlying mechanics, in relation to FEP and memory,
463 the performance of computational learning within the EAP
464 hydrogel can be assessed using the same Pong game-based ac-
465 tivity, in a simulated game environment, utilising the reservoir
466 computing framework.

Experimental Design. As in the paper "In vitro neurons learn
468 and exhibit sentience when embodied in a simulated game-
469 world" (21) a MEA is used to interface the EAP hydrogel with
470 a computer system containing the Pong game environment.
471 These electrodes are used to provide stimulus as input to the
472 gel via the application of electric fields at select locations, and
473 record the gel's distribution of ions as output through sensing
474 the current draw at select locations. A region of the gel was
475 designated for stimulation (receiving game environment infor-
476 mation).

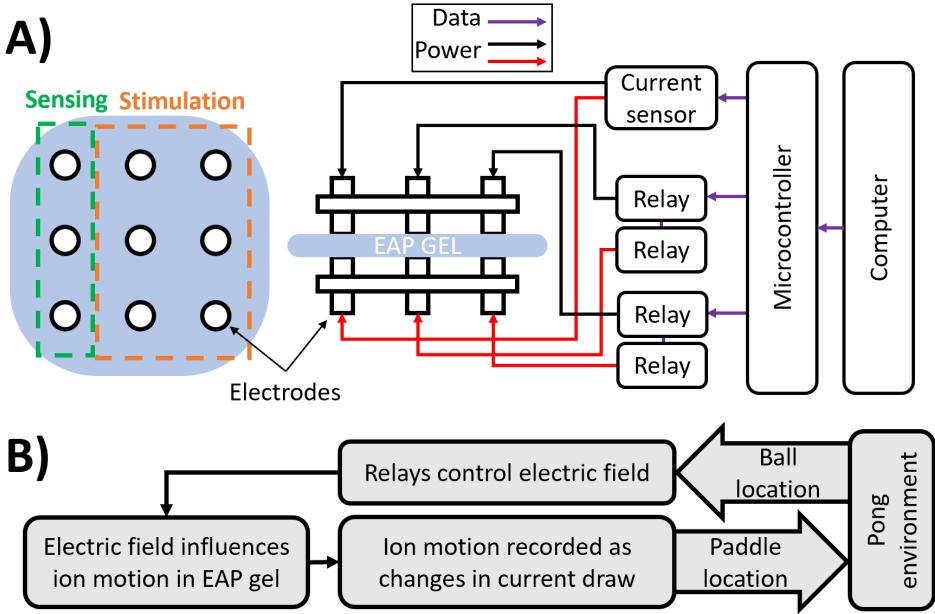


Fig. 3. System layout of closed loop communication between the computer containing the Pong environment and the EAP hydrogel. **A)** The layout of communication between components in the system and separation of the electrodes into regions. The electrodes are divided into driving, which provide stimulation, and sensing, which measure current draw. The driving electrodes are driven by relays that direct the electric field (20 V). The sensing electrodes provide a small voltage (2 V) and the current draw is measured by a current sensor. Both relays and sensors are coordinated by a microcontroller, which is in turn directed by the computer and Pong game environment. A more detailed circuitry schematic can also be found in the supplementary information figure S1, additionally an image of the physical hardware can be seen in the supplementary information figure S2. **B)** A flow chart of the information path from the Pong environment on the computer to the hydrogel containing the ions.

mation as the location of the ball), acting as the stimulation layer of the reservoir framework. Another region of the gel was designated for sensing (providing information used as motor commands for the paddle) acting as the readout layer of the reservoir framework. This approach is similar to how regions are designated in MEA applications (21, 55). The layout of these regions can be seen in figure 3 section A.

Electro-mechanical relays were used to apply the electric field as they allow for a high impedance state where electrode pairs are not connected to either power or ground, this prevents electric field interference from electrode pairs that are not in use. Current sensors were used to measure the current draw at the sensing electrodes. These relays and sensors were coordinated by a microcontroller that sends and receives commands from the computer containing the Pong environment. The layout of these components is illustrated in figure 3 section A. A more detailed circuitry schematic can also be found in the supplementary information figure S1, additionally an image of the physical hardware can be seen in the supplementary information figure S2. Through this layout of hardware and software, a closed loop was established where the Pong game's environment, represented as the ball location, was presented to the gel as a shaped electric field. Ion concentrations were then provided to the Pong game to move the paddle. This closed loop is outlined in figure 3 section B. The computer sends the current stimulation instructions to the MEA, on receiving this the MEA sends back the recorded current values at that time. Due to the call and response setup the computer governs the iteration rate of the closed loop, in this experiment this is set to 0.5 seconds per loop.

The location of the ball in the game environment is sent from the computer to the MEA, on each closed loop iteration, as a binary string where each position represents 1 of 6 regions.

The Pong game environment is divided into 6 regions, shown in figure 4 section B, as the ball passes through those regions the corresponding electrode pair is activated to stimulate that area. The regions as they are labelled in black in figure 4 section B are how they will be referenced throughout this paper. The MEA interprets the binary string into instructions for each region, activating the corresponding relays and applying the electric field accordingly. The ball is wide enough to be in multiple regions at once. As the ball passes from one region to the next both will be briefly active, these combinations allow for more states to represent ball position. The game environment can be seen on the computer screen along with the apparatus in the supplemental information figure S3.

The current sensors record the current in mA, sending that information to the computer, as shown in figure 4 section A. As there can be small inconsistencies between gels the starting current value is measured, before the Pong game is started, using the average current value over 75 seconds. This value is used as a baseline to calibrate the sensor readings. During the closed loop when the current values are received by the computer from the MEA, in response to the stimulation instructions, they are plotted on the display, as shown in figure 4 section A. After being plotted the values are normalised using the maximum and minimum current values recorded from the hydrogel during earlier experimentation with the apparatus, found to be +3mA and -0.75mA respectively. A trend is then calculated, using these normalized values against their locations represented in the Pong game environment; black for top, brown for middle, and red for bottom. The trend is generated by fitting a 2nd degree polynomial to the 3 points. The peak of this trend is the predicted point of highest current, and highest ion concentration, in the gel and is where the paddle is placed. This is illustrated in the example in

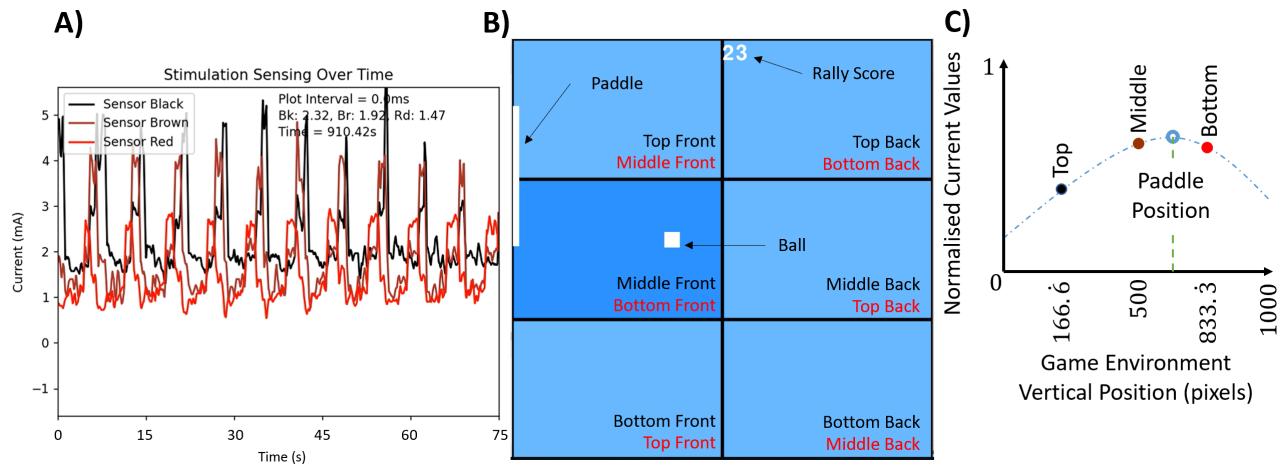


Fig. 4. The software representation of the Pong game within the computer. **A)** The current sensor readings as they are received by the computer. The black, brown, and red lines represent the top, middle, and bottom sensors respectively. The interval between received data is also measures and averages at 0.5 mS, labelled as "Plot Interval". The averaged current across the graph for each sensor is measured for the purposes of establishing default values, labelled as "Bk", "Br", and "Rd". The total recording time is also displayed labelled as "Time". These sensor values move paddle in the Pong game environment through conversion to a position based on sensor location, as visualised in figure 4 section C. This can be seen on the computer screen along with the apparatus in the supplemental information figure S3. **B)** The simulated Pong game environment, which is separated into 6 regions as described in figure 3 section B. When the ball is in a region it will darken and the corresponding electrode pair will stimulate the EAP hydrogel. The score is displayed at the top and resets when the ball hits the left wall behind the paddle. The regions are also labelled in black here for reference within this paper. The paddle is set to 1/3 the environment height. The rearrangement of regions used for the baseline comparison dataset, where 'vision' is impaired, is indicated by the red region labels. This can be seen on the computer screen along with the supplemental information figure S3. **C)** An example of the method used to place the paddle based off the current readings as measured in figure 4 section A, shown here as the same representative colours and labelled with their region locations. The current readings are normalised between 0 and 1 and plotted based on their region's centre vertical location. A 2nd degree polynomial is applied, as indicated by the blue dashed line. The point where the polynomial is at its peak is the maximum predicted current, as indicated by the blue circle, and this is where the paddle is placed.

543 figure 4 section C.

544 At the start of the Pong game the ball starts in the centre of
 545 the game area with a random speed and direction towards the
 546 right wall, between 4 and 8 pixels per game update step for the
 547 x axis and between 1 and 8 pixels per game update step for the
 548 y axis. This allows the ability of the hydrogel to be assessed
 549 against various ball trajectories and speeds giving a more
 550 accurate measurement of performance. When the ball hits a
 551 surface, such as a wall or paddle, the speed perpendicular to
 552 the surface is reversed to allow it to bounce with approximate
 553 realism and allow the ball path to be predictable. If the ball
 554 hits the paddle the score is increased by 1, if the paddle misses
 555 the ball and the ball hits the wall behind the score resets to
 556 0 and the ball resets position to the middle of the play area
 557 with a new random speed and trajectory.

558 The top and bottom regions are equally hit by the ball twice
 559 as often as the middle section. This is expected in the course
 560 of a Pong game as the ball most often has a diagonal trajectory,
 561 and the top and bottom regions are the boundaries. This hit
 562 distribution was recorded and is shown in the supplementary
 563 information figure S4 to show no bias in ball hit locations. The
 564 full experimental procedure for each run of the Pong game
 565 with the EAP hydrogel can be found in the supplementary
 566 information section "Pong Game Experimental Procedure".
 567 In total 21 separate EAP hydrogel runs were carried out,
 568 collecting 3500 seconds of 'game play' for each run before the
 569 hydrogel degraded beyond the point of continuing. All data
 570 from these separate runs was combined into a single dataset to
 571 view the overall averaged behaviour of the hydrogel, and assess
 572 the performance as well as the repeatability. Once ran and
 573 analysed the results were deemed stable enough to not require
 574 additional repetitions as data standard deviation was stable,
 575 more detail on this can be found in the following section.

576 A set of experiments were also carried out to create a
 577 baseline to compare performance against. The goal of this
 578 experiment was to explore the null hypothesis, that the per-
 579 formance increase observed in the hydrogel was not a result of
 580 the accurate environmental information the gel received. To
 581 explore this the simulated environment regions correlation to
 582 the gel's stimulation regions were rearranged, so that the gel
 583 received incorrect information about the balls position. The
 584 rearrangement of game environment regions can be seen in
 585 figure 4 section B, indicated by the red region labels. For
 586 the null hypothesis to be rejected the hydrogel would need
 587 to perform worse than if the paddle was moving randomly,
 588 as the paddle is 1/3 the height of the simulated environment
 589 the rejection would require a hit/miss ratio below 33%. This
 590 experiment was run 10 times to assess the performance and
 591 repeatability. Once ran and analysed the results were deemed
 592 stable enough to not require additional repetitions as all data
 593 recorded satisfied the rejection of the null hypothesis, more
 594 detail on this can be found in the following section.

595 **Results.** To assess the performance of the hydrogel within the
 596 Pong game, and to discern the extent of any learning achieved,
 597 the ability of the hydrogel to hit the ball must be quantified.
 598 To this effect the hit to miss ratio was calculated, using the
 599 combined dataset of the 21 game runs, and plotted against
 600 time. The data is segmented based on the region in which
 601 the hit/miss occurred. Due to the physics of the ball's motion
 602 the top and bottom regions hit more often than the centre,
 603 this can be seen in the hit distribution in the supplementary
 604 information figure S4. Because of this it is also possible that
 605 some regions experience more learning than others, due to the
 606 frequency of the ball being in that region. By segmenting the
 607 data based on the region of the hit, the performance over time

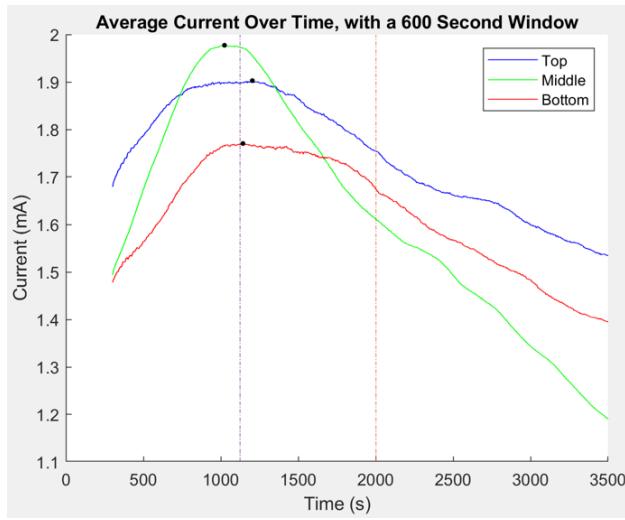


Fig. 5. The current measured at each sensor over time for the combined unpaired dataset of runs, smoothed using a moving average window of 600 seconds. Because of this moving average the data starts at 300 seconds, the mid point of the first averaged dataset subsample. The point of maximum current draw was found for each data trend and marked with a black dot. 1203, 1023, and 1142 seconds for top, middle, and bottom respectively. The average point of maximum current draw is marked by the purple vertical dashed line at 1122 seconds. The orange vertical dashed line marks the earliest point that the EAP gels broke down. Standard error was calculated for each trend from the windowed samples and found to be at maximum, 0.0169, 0.0053, and 0.0116 for the top, middle, and bottom trends respectively.

can be shown without being overshadowed by the performance of other. A moving window was used to find the hit/miss ratio over the course of the hydrogels playthrough of the game. A window of 600 seconds was used, this window size provided the clearest view of the performance within the game and is used for all further data analysis, additional windows sizes can be seen in the supplementary information figure S5. Because of this moving average all data smoothed in this way starts at 300 seconds, the mid point of the first averaged dataset subsample.

To check the stability of results from the 21 experimental repetitions, the stabilisation of the standard deviation in the recorded hit/miss ratio of the dataset was analysed as more experimental results are added to the dataset. The standard deviation was applied to datasets over time, consisting of varied numbers of experimental runs (1 to 21), using the same 600 second window. As the dataset becomes larger the changing standard deviation becomes more stable and undergoing less change between datasets. With a dataset consisting of 19 experimental runs the trend of standard deviation does not appear to change any further and so has stabilised and no further experimental runs are necessary. The plots of these standard deviations over time with varying dataset sizes can be seen in the supplemental information figure S6.

Although 3500 seconds of play time was recorded, the EAP hydrogel's polymer structure breaks down after prolonged stimulation. With extended application of the electric field the hydrogel undergoes electrolysis due to contact with the electrodes. This electrolysis breaks down the polymer structure (56) as well as oxidising the electrodes causing rust to form, interfering with the ionic concentrations within the gel. This breakdown of the hydrogel impacts its functionality and disrupts connections between the sensing and stimulation

electrodes by limiting ionic migration. This breakdown in structure can be seen visually as large holes form in the material, physically preventing further migration of ions, shown in the supplementary information figure S7.

The data beyond the point of degradation does not represent the true mechanics of the EAP hydrogel, so a cutoff point was selected after which the hydrogel was too degraded for results to be representative of the performance. This cutoff point was selected by analysing the current drawn at each sensing electrode over the course of the combined dataset, a figure of this data can be seen in figure 5. The figure represents the combined data of all 21 runs averaged using a 600 second window as described previously. The current initially rises as ions move towards the electrodes. As the ions collect their rate of migration decreases due to the increasing concentration gradient. Eventually the ions stop moving, the current reaches its maximum, and the gel breaks down. This can be seen in figure 5 as the current reaching its peak, indicated by the vertical purple dashed line, then gradual falling as the gel breaks down. As the gel breaks down further, the current continues to reduce as the connection between the sensing and stimulation electrodes, created by the polymer structure, is broken down. The earliest point at which holes were observed to form within the hydrogel, as shown in the supplementary information figure S7, is marked in figure 5 by the orange vertical dashed line at 2000 seconds. This is just before the point in the experiment where the measured current is lower than the initial starting current, indicating the hydrogel is now less conductive than it when it started and incapable of any further performance. Because the data beyond 2000 seconds represents a system too degraded to perform, the graphs used to analyse the hydrogel's performance are limited to between 0 and 2000 seconds so to remove any unrelated data.

Using these analysis techniques, the hit/miss ratio graph was generated and can be seen in figure 6 section A. From this plot a clear learning curve can be seen in the performance of the paddle when in both the middle and bottom regions of the game environment. Each region has a unique starting hit/miss ratio.

For the top region the paddle rarely missed the ball. However, the boundary conditions are considerably easier to hit as the trend calculated from the sensed currents at those regions is much simpler. Additionally, the top sensor tended to read higher currents than the other regions even after calibration, this is likely due to some small differences between the distances in electrodes in the MEA rig. This was however consistent through all experiments, so did not impede the performance analysis.

For the bottom region the paddle initially hit the ball approximately four times more often than it missed, hitting 79% of the time. As the bottom region is also a boundary condition this is expected. However, over time this performance did increase up to a maximum of 87% hit rate, showing an improvement of 8% over 1750 seconds.

For the middle region the paddle had an initial ball hit rate of 50%, but over the course of the game this rose to maximum of 60%, giving a rise of 10% over 1450 seconds. The middle region shows the greatest improvement in performance, this makes sense as the middle region had the lowest initial performance leaving more room for improvement. Improvement in performance requires more effort for each unit increase, for

example 20% to 30% requiring more effort than 10% to 20%. This diminishing return in improved performance can be seen in machine learning and game theory (58, 59).

To better show the overall performance increase through the course of the game, without the individual regions overshadowing improvement, a graph was generated of the rally lengths against time. When the paddle hits the ball, the score increases and when the paddle misses the score resets to zero. The score reached before it resets to zero when missed is the rally length. Each time the ball was missed the rally length achieved was recorded against the time, and plotted using an averaging window of 600 seconds as with the previous plots. This plot is shown in figure 6 section B, and shows an increase in rally length over the course of the game. This graph clearly shows the ability of the gel to perform within the Pong game before and after the learning period. With this separation of states, before and after learning, the significance of the change in the hydrogel's performance can be analysed through a statistical significance test, using $\alpha = 0.05$ as the null hypothesis threshold, between the rally length before and after the performance increase or "learning". First the samples, from the rally length dataset, representing before and after the learning need to be defined, through the trend of the change in rally length the point at which learning starts and the point at which maximum learning occurs can be defined. To this end a 4th degree polynomial best fit line was applied to the data, shown in figure 6 section B as the blue dashed line. This was found to be the lowest degree polynomial that still visibly fit the data trend. Using the best fit line, the average rally lengths started around 4.4 and over the course of the game increased to 5.1, giving an increase of 0.7 over the course of 1700 seconds. With this trend, the samples can be extracted from the rally length dataset prior to smoothing.

The sample representing the learned state can be extracted centred on the point of maximum learning, marked in figure 6 section B as the black dot, at 1744 seconds. To ensure the largest sample possible the extracted data will be from the range 1487, represented in figure 6 section B by the purple dashed line, to 2000 seconds giving 228 data points. The sample representing the pre-learned state can be taken from 0 seconds up to the point where performance starts to improve at 1000 seconds, as evident by the trend in figure 6 section B. To accurately represent the pre-learned state the sample must span the largest range, however to accurately compare the samples they must be the same size. To achieve this the full data range from 0 to 2000 seconds was sub-sampled into 228 data points, making the range of the sample 0 to 838 seconds. Although the pre and post learned values were recorded from the same gel samples, many data points were recorded per sample. Therefore, the variables representing before and after learning can be considered independent, as there is no one-to-one relationship between data points in the two states.

due to the direction of the one-tail distribution, and σ_U is the standard deviation of U . The Mann–Whitney U test can be used to assess significance as both samples have non-normal distributions and independent variables. This is evident when the samples are represented via a density histogram as shown in the supplementary information figure S8 where both follow an exponential distribution due to 0 being the most common rally length in both samples.

Following the Mann–Whitney U test, the ranked sum was calculated for each sample, where ties were present the assigned rank was equal to the midpoint of unadjusted rankings. The summed ranks were calculated to be 47476.5 and 56719.5 for the pre and post learned samples respectively. The lower of the values was used to calculate the U statistic using equation 8 (61), where U is the U statistic, R_s is the lowest ranked sum value, and n is the size of the sample which here is equal in both samples. σ_U is calculated using equation 10 (62), where C_{ties} is the adjustment due to the number of tied ranks. C_{ties} is calculated in equation 11 (62), where t_k is the number of ties for the k^{th} rank, and K is the total number of unique ranks with ties.

$$p = \Phi(z) \quad [6]$$

where:

$$z = \frac{U - \mu_U + C_{cont}}{\sigma_U} \quad [7]$$

$$U = R_s - \frac{n(n+1)}{2} \quad [8]$$

$$\mu_U = \frac{n^2}{2} \quad [9]$$

$$\sigma_U = \sqrt{\frac{n^2}{12}(2n - C_{ties})} \quad [10]$$

$$C_{ties} = \frac{\sum_{k=1}^K (t_k^3 - t_k)}{2n(2n-1)} \quad [11]$$

These calculations give a $z = -3.3433$ which, using equation 6, gives a $p = 0.00041$. Using the z table for the negative score this places the probability below 0.05% and rejects the null hypothesis, validating the improvement before and after learning is statistically significant. This reinforces the observed increase in performance of the hydrogel in the Pong game environment, by showing that the improved hit rate was not just limited to the regions in which they were associated, but contributed to improved performance in the game as a whole.

It can also be observed that the peak in rally length observed in figure 6 section B, aligns with the average peak in hit rate observed in figure 6 section A, with a slight delay of around 150 seconds. This delay makes sense as the effect of the improved hit rate would take some time to translate to an improvement in rally length, due to the additional layer of abstraction caused by the time taken for a full rally to be completed. Further to this, the average peak in hit rate observed in figure 6 section A, aligns with the peak current measurement observed in figure 5, with another delay of around 600 seconds. As the ions move to their final location, representing the environment, the current increases. This increase in current represents the energy landscape of the hydrogel being influenced by the received environmental data, and filling

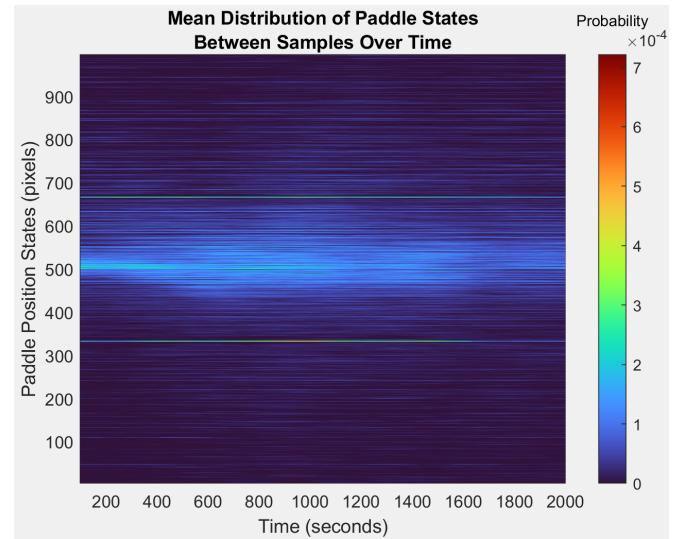


Fig. 7. The mean distribution of paddle states between all 21 experimental samples, displayed as a temperature plot. The distributions were sampled every 20 seconds starting at $t=100$ using a window of 200 seconds, with window size provided enough datapoints for accurate representation with minimal overlap between samples. The temperature of each point represents the probability of that state. The states of 0 and 1000 pixels have been omitted, the probability of these states was much higher than others (0.0129, and 0.0177 respectively) due to the boundary conditions and so obscured the rest of the data by biasing the scale. An additional set of plots can be seen in the supplemental information figure S9, showing individual distributions for timestamps 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, and 2000 using a larger window for direct comparisons of sample distributions.

the memory similarly to how the memory component was observed in the first section of this paper, via the current measurement. As with the delay between hit rate and rally length, it would take some time for the base level current measurements to translate to improved hitting ability, due to the layers of abstraction caused by the mechanics of the game and MEA interface.

Although these results show the increase in performance and, through the current plot in figure 5, link this increase in performance to the memory mechanics and ion motion, these results do not give much insight into the layers of abstraction between ion motion and improved hit/miss ratio. In other words, how does increased performance present itself in paddle behaviour respect to ball position throughout the game, and how does this behaviour link to the minimisation of free energy via ion migration. To answer this question, and further develop the chain of events that culminate in improved performance, the behaviour of the paddle can be analysed through the course of the game. This can be approached through analysing the paddle's motion in response to the ball's location, or through analysis of the paddles positional distribution and how it changes over time related to the ball's positional distribution. The latter provides a more in depth view of the layers of abstraction and their relationship to the base mechanics at play and will be pursued in this study. However, the former is included in the supplementary information under the heading "Paddle Motion and Standard Deviation Analysis".

The paddle is limited to move in the y-axis, each vertical position the paddle exists in can be thought of as a state that the paddle inhabits. By applying a window to the data, as with previous plots, the change in state distributions through

the course of the game can be analysed. Figure 7 shows the changing mean paddle state probability distributions over the course of the pong game. Individual distributions can be seen in the supplemental information figure S9, showing probability distributions for timestamps 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, and 2000 using a larger window for direct comparisons of sample distributions. With these distributions there is a clear change in shape as the game progresses, initially starting with a large peak in the center that becomes less pronounced as the game reaches its maximum performance point at 1800 seconds, as found from figure 6 section B, where the distribution becomes more uniform. From this it can be deduced that the distribution of paddle states changes through gameplay, by quantifying this change the mechanics involved can be explored and linked to the initial theories of free energy minimisation. In figure 7 the states of $y=0$ and $y=1000$ were omitted as their larger values obscured the rest of the results. These larger values are a result of the boundary conditions, to fully explore the learning mechanics the impact of these boundary conditions must be investigated.

The boundary conditions arise as the paddle cannot pass beyond the edge of the game environment. If the trend in current used to place the paddle, as described in figure 4 section C, has a positive gradient toward either edge the paddle will be placed at the edge, this means that there are more combinations of current that lead to the paddle being at the edges than in-between. The boundary conditions can be visualised by simulating the paddle with random current inputs, using the algorithm described in figure 4 section C, and recording the paddle's positional states. A distribution of these simulated results is shown in figure 8 section A, this plot shows that with random current the paddle has a significant tendency toward either end with a peak distribution in the centre. The shape of distribution at the centre can also be explained as a result of the boundary conditions and is analogous with a depletion force. The boundaries in the paddle's motion act as two point forces with high attraction to the paddle, the forces interact at the mid-point between them creating a peak (63, 64). With greater resolution the effect of the boundary conditions could be reduced, however that is beyond the scope of this study. The distribution shown in figure 8 section A is similar in shape to the initial distribution at $t=100$ in figure 7, as the paddle's initial motion is almost entirely influenced by the boundary conditions having yet to gather enough information to diverge.

Although the boundary conditions explain the initial state distribution of the paddle, to understand the change in behaviour the distribution at maximum performance ($t=1800$) must be assessed. In an ideal pong game the paddle position matches that of the ball's vertical position at all times, meaning that in an ideal game the ball and paddle should have identical distributions of vertical positional states. By comparing the paddle's state distribution against that of the ball, the paddle's change in behaviour can be assessed against what would be the ideal behaviour, this allows analysis of what attributes of the paddle's distribution change to match that of the ball leading to improved performance. The ball's distribution can be simulated using the pong game environment. However, the ball resets on missing the paddle which alters the ball's distribution, this means that as performance improves the ball's behaviour changes as it is hit more and

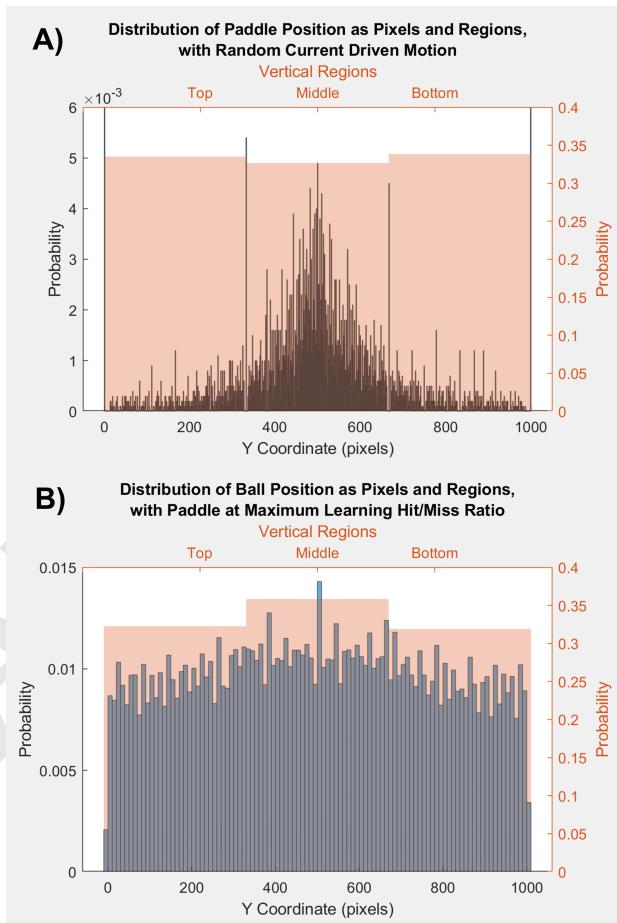


Fig. 8. Distributions of positional states from paddle and ball simulations. Each plot shows the distribution divided into bins of each state, signified by the blue bars, and divided by regions shown by the orange transparent overlay bars. **(A)** The simulated distribution of paddle states generated, using the algorithm described in figure 4 section C, with randomised current values. 10000 samples were simulated to show the distribution. To allow the central peak to be visible the left y-axis has been limited to 0.006, however the states of 0 and 1000 are higher than the central peak due to the boundary conditions. The states 0 and 1000 have probabilities of 0.2754 and 0.2796 respectively. **(B)** The simulated distribution of vertical ball states using the maximum hit/miss ratio for each region. The ball was simulated with the same game physics, starting with a random direction and resetting to the centre on a miss. When the ball hits the paddle wall the decision of if it will bounce follows the ratio at maximum performance, 0.98, 0.58, and 0.86 for the top, middle, and bottom regions respectively. As the ball moves the full path of each ball simulation was recorded resulting in 600000 samples, with the travel of each ball simulation before reset lasting on average 245 samples.

follows different trajectories. This is evident when comparing two simulations of the ball's positional distribution, where the ball is never hit by the paddle and where it is always hit by the paddle, shown in supplementary figure S12. To accurately compare the paddle's distribution to that of the ball, the ball's distribution must represent the behaviour for the ball at the point of maximum performance. At this point the hit/miss ratios are 0.98, 0.58, and 0.86 for the top, middle, and bottom regions respectively, shown in figure 6 section A. When the ball hits the paddle wall the ball is bounced according to the region's hit/miss ratio, this distribution can be seen in figure 8 section B. From this distribution the ball's vertical motion over the game environment is more uniform than the paddle's initial distribution with fewer unique states, indicative of the predictable ball trajectories. From figure 7, the paddle's distribution appears to become more uniform as the central peak becomes less pronounced, additionally the number of unique states reduces, also moving closer to the ball's distribution. These observations however need to be quantified via metrics to properly present the continuous behavioural change of the paddle's motion, and the change in the ball's motion as a result.

There are two main metrics used to analyse the attributes of the paddle and ball's motion distribution, both representative of entropy within the system aligning with theories of FEP and Bayesian inference (22, 25, 32–34).

- Number of unique states: The number of unique states/positions the paddle/ball is recorded in within the sample window used to generate the distribution. Defined in equation 13 as o , where Q is the sample used to make the distribution, O is the set of elements in Q without repetitions, and $|O|$ represents the cardinality of O . When applied to the ball, due to the resolution of the ball's motion as applied to the gel, the number of unique states of the ball will remain constant, limited by the 6 regions the ball inhabits, so not providing usable information. Unique states is representative of entropy as the paddle position is determined by conductivity of ions and their distribution, in statistical thermodynamics entropy is defined as $S = k_B \ln \Omega$ (65). S is the entropy, k_B is the Boltzmann constant, and Ω is the number of microstates which in this case is analogous with the number of unique states u .

- Variance in frequency of unique states: The variance in quantity of occurrence of each unique state, used as a measure for uniformity in the distribution. Defined in equation set 12 as σ^2 , where c_i is the number of occurrences of the i^{th} element of O , \bar{c}_{arc} is the mean value of c , Q_j is the j^{th} element of Q , and q is the number of elements in Q . Shown in figure 8 section A, the boundary states of 0 and 1000 greatly overshadow the central peak distribution. As the edge conditions are results of the interfacing algorithm and only represent a fraction of the paddle's motion, they can be ignored for the purposes of calculating variance in the paddle. This allows the uniformity in the central peak to be more accurately compared to that of the ball which is not subject to boundary conditions, as evident by figure 8 section B. Variance is linked to entropy through information theory, although not directly proportional entropy and variance are both

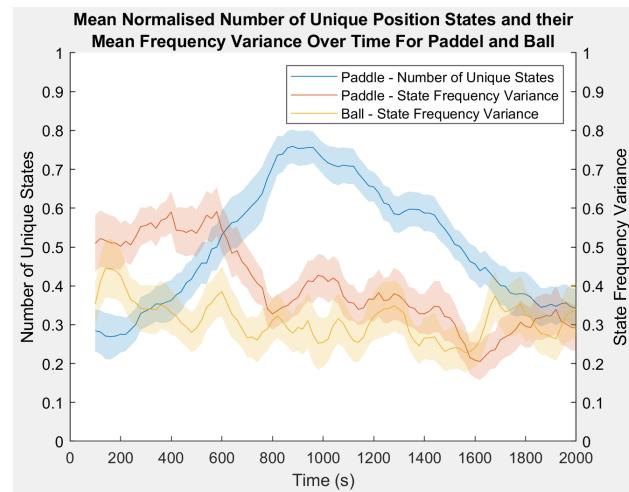


Fig. 9. The mean of the normalised number of unique paddle states and the mean of the normalised variance between unique paddle state frequencies for both the paddle and ball, over the course of the game. Distributions were generated from each gel sample using a window size of 200 seconds, as in the plots in figure 7, sampled every 20 seconds from $t=100$ to $t=2000$. The number of unique paddle states and state frequency variance were then normalised and the mean between gel samples found. The standard error between gel samples is show by the shaded area.

measures of expected values and share trend directionality (66).

$$\sigma^2 = \frac{\sum (c_i - \bar{c})^2}{o - 1} \quad [12]$$

$$o = |O| \quad [13]$$

Where:

$$c_i = \sum_{j=1}^q f(Q_j, O_i)$$

$$f(Q_j, O_i) = \begin{cases} 1 & Q_j = O_i \\ 0 & Q_j \neq O_i \end{cases}$$

Due to the dynamic nature of the hydrogels, each sample will have slight variations in ion distribution, polymer density, and surface texture, this could lead to differences in recorded metrics. However, the learning behaviour is still driven by the same mechanics regardless of starting inconsistencies, and as a result the trend shape of these metrics will be consistent between samples. To minimise the affect of gel synthesis inconsistencies, the metrics are normalised via feature scaling to better highlight the shape of the trends. The plot of these trends is shown in figure 9, the lines represent the mean between all samples with error bars representing the standard error between all samples used. This plot shows several interesting features when compared to previous plots.

1. At $t=200$: The number of unique paddle states is low while the variance in both paddle and ball is high, showing few states with large variations in frequency. This matches the simulated distribution form figure 8 section A, expected as the initial distribution will be almost entirely driven by the boundary conditions, as no information has been gathered about the game environment and ions remain in a homogeneous distribution.

- 982 2. At $t=600$: The number of states increases as the variance
 983 decreases in both paddle and ball, this happens as more
 984 information about the environment is gathered and co-
 985 incides with the rise in current seen in figure 5. With
 986 the application of an electric field the ions start to mi-
 987 grate within the hydrogel to represent the stimulation,
 988 as ion mobility increases the number of unique states
 989 increases and the paddle explores more of its working
 990 area. Simultaneously as more of the working area is ex-
 991 plored, paddle motion becomes more evenly distributed
 992 and so the variance decreases. As the ball is hit more
 993 often its variance decreases as it moves through more
 994 of the game environment before being reset on a miss,
 995 evident when comparing variance of simulations where
 996 the paddle always hits and always misses the ball shown
 997 in supplementary figure S12.
- 998 3. At $t=1000$: The number of states reaches its maximum
 999 and starts to decrease, this coincides with the behaviour
 1000 of the current in figure 5 and can be explained though ion
 1001 mobility. Initially the ions have total mobility as the gel is
 1002 yet to change in structure, as the ions mobility increases
 1003 hysteresis takes affect within the hydrogel opposing ion
 1004 migration. Eventually ion mobility reaches its maximum
 1005 causing the maximum peak in current seen in figure 5,
 1006 after this point the hysteresis effect continues and ion
 1007 mobility decreases causing the number of unique states to
 1008 decrease, as ions start to settle into their final positions.
- 1009 4. At $t=1800$: Both the number of states and variance in
 1010 the paddle have reached their minimum, this coincides
 1011 with the point of maximum performance shown in figure
 1012 6 section A and B. The normalised variance in paddle and
 1013 ball now match as the paddle and ball's behaviour have a
 1014 stronger coupling. As the memory continues to saturate
 1015 the ion distribution comes to better represent the ball's
 1016 behaviour, this leads to reduced paddle state variance
 1017 and reduced number of unique paddles states, causing the
 1018 behaviour of the paddle to become more like that of the
 1019 ball. This causes the ball to be missed less often, resulting
 1020 in a more homogeneous distribution of ball states and less
 1021 variance, as seen in supplementary figure S12 section B.
 1022 The saturation of memory is cause by the hysteresis and
 1023 so is linked to the breakdown of the polymer structure,
 1024 eventually the same mechanic that allows the gel to retain
 1025 memory causes the polymer structure to break down to
 1026 a point that it inhibits conductivity and performance
 1027 reduces.

1028 The trends shown in figure 9 and the way in which they
 1029 coincide with features in other figure adds another step to the
 1030 chain of events that lead to the increased performance, and
 1031 leads to interesting forms of emergent behavior occurring as
 1032 the ions migrate to reduce free energy.

1033 Baseline results were collected to further show that the
 1034 hydrogel's improved performance was a direct result of the
 1035 environmental information received, and explore the null hy-
 1036 pothesis that the improved performance is simple cause and
 1037 effect within the gel and thus not a result of learning. As de-
 1038 scribed in the experimental description the electrode positions
 1039 were rearranged to give incoherent environmental information
 1040 to the hydrogel, effectively impairing its 'vision'. The hit/miss

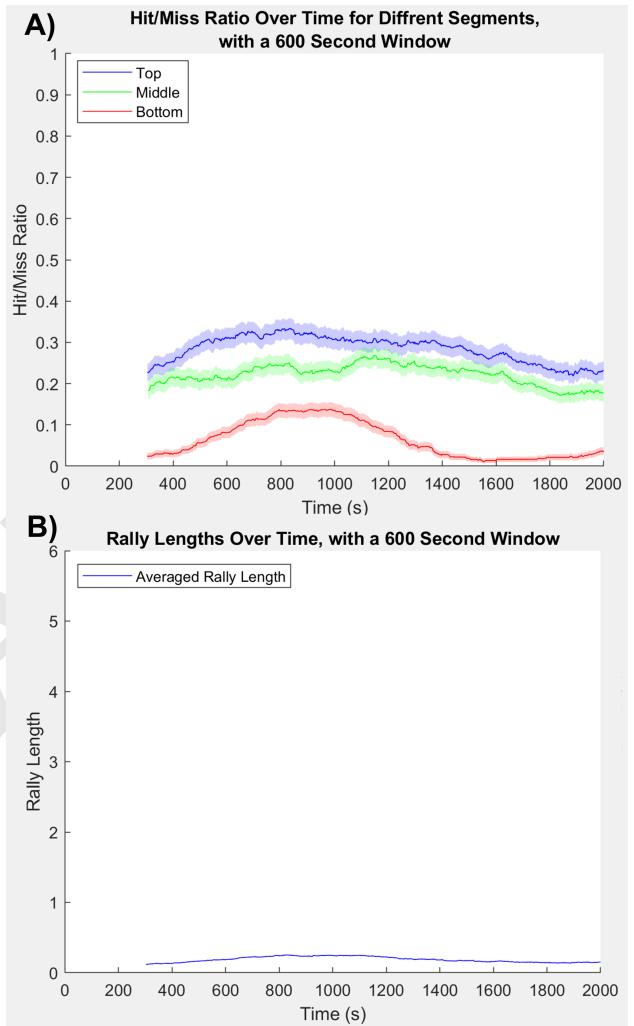


Fig. 10. A) The hit to miss ratio over time for the combined impaired feedback loop dataset, separated based on the region the hit/miss occurred in as described by figure 4 sections B. Smoothed using a moving average window of 600 seconds, as with previous figures. The standard error for each windowed sample is show by the shaded area. The performance in the Pong game is very poor, not reaching much beyond 0.3 for any of the regions. **B)** The rally length (maximum score before a miss) against time for the combined impaired feedback loop dataset. Smoothed using a moving average window of 600 seconds, as with previous figures. The standard error was calculated for each windowed sample, with a maximum of 0.0259. The Pong game was barely able to reach a rally of more than 1 point through the entire dataset.

ratio was recorded in the same way, shown in figure 10 section A, using the same averaging window and cutoff point as with the previous plots. Additionally, as with the previous experiment, the stabilisation of the standard deviation in the recorded hit/miss ratio of the dataset was analysed as more experimental results are added to the dataset. The standard deviation was applied to datasets over time, consisting of varied numbers of experimental runs (1 to 10), using the same 600 second window. As the dataset becomes larger the changing standard deviation becomes more stable and undergoing less change between datasets. With a dataset consisting of 9 experimental runs the trend of standard deviation does not appear to change any further and so has stabilised and no further experimental runs are necessary. The plots of these standard deviations over time with varying dataset sizes can be seen in the supplemental information figure S11.

From the plots in figure 10 section A little improvement in performance can be seen, with no region achieving a hit ratio beyond 30%. The middle region that showed the most improvement in the previous results shows little to no change in performance over the course of the game. Similarly, the top and bottom regions show some initial improvement but far too little to be considered significant especially given the low initial hit ratios. This lack of performance is even more evident when the rally lengths are plotted, shown in figure 10 section B. From this graph little to no improvement can be seen in performance, with the average rally length well below 1 point. This means that the ball was missed far more often than hit throughout the game, with no consecutive hits achieved.

Due to the hit/miss ration not reaching above 33% the null hypothesis can be rejected as the performance was worse than if the paddle were moving randomly, given the paddle length is a third the total game environment height. Even with this incorrect environmental information the hydrogel is still being provided stimulation, this means the hydrogel is still learning as the ions move to a distribution representative of the information provided. Unfortunately, as the information provided is wrong, what is learnt by the hydrogel is also wrong, leading to its performance being worse than random motion. This observation helps to reinforce that the hydrogel's increase in performance is directly related to it being presented with an accurate representation of the virtual environment in which it is acting.

Discussion. From the results observed here there is a clear increase in performance of the hydrogel to play the Pong game, as shown in figure 6, when given information that is representative of the virtual game environment. The development of the increased performance can also be directly linked to the underlying mechanics of the EAP hydrogel through ion migration and memory mechanics due to hysteresis. Through this linking of mechanics a logical similarity can be established between the properties of EAP hydrogels and theories of learning via FEP.

In FEP learning, external states of the environment influence internal states of the agent, in this case the EAP hydrogel, through a looped exchange of information called active inference. Active inference analogises an internal generative sudo model used to predict inputs that represent the external world (22, 25, 32), based on theories of Bayesian inference (33, 34). This feedback of information flow is illustrated in figure 11

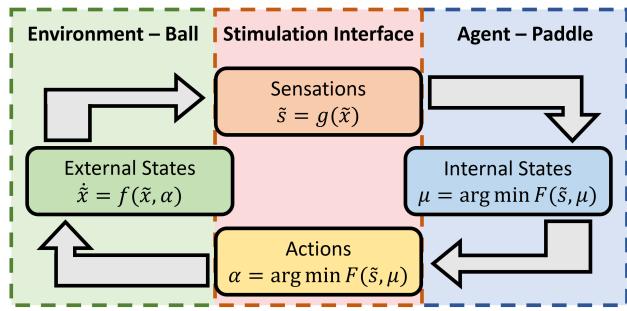


Fig. 11. Diagram of information flow and dependency between states of the free \tilde{s} energy minimisation task derived from FEP equations (22, 67, 68). External states \tilde{x} are of the environment in which the agent acts, synonymous with the probability distribution of the ball. Sensations \tilde{s} are the interaction by which the agent gathers information, synonymous with the regional stimulation. Internal states μ are of the agent's internal model, synonymous with the probability distribution of the paddle. Actions α are the interactions between agent and environment, synonymous with sensed current interpretation.

with equations showing informational exchange. Environmental states \tilde{x} are described by the constraints of the game physics defining the motion of the ball. \tilde{x} translate to sensations \tilde{s} via the agent's interface, in this case the stimulation interface translates the ball's pixel positions into regional stimulations. As such \tilde{s} is synonymous with the probability distribution of ball regions represented in figure 8 section B, supplemental figure S12, and used for state frequency variance of the ball in figure 9. \tilde{s} translate to internal states μ in the agent as a minimisation of the agent's internal free energy F , the EAP hydrogel's internal structure is representative of these internal states, synonymous with the sudo internal model analogy of FEP learning and updated as sensations of information are received. μ causes actions α via the minimisation of F and as such μ is synonymous with the probability distribution of paddle states represented in figure 8 section A, and used for number of unique states and state frequency variance of the paddle in figure 9.

It is clear from 9 the minimisation of the unique number of states and state frequency variance in both paddle and ball coincide with improved performance, and through figure 11 their part in the feedback of information can be seen as \tilde{s} and μ . This link can be further highlighted through formal definitions of free energy F bound on Bayesian model evidence. F is defined in equation 14 (67, 68), where D_{kl} is the Kullback–Leibler divergence, $Q(\mu)$ is the probability density of internal states μ , and the joint probability distribution of $P(\tilde{s}, \mu)$ is such that $P(\tilde{s}, \mu) = P(\tilde{s}|\mu)P(\mu)$.

$$F(\tilde{s}, \mu) = \underbrace{D_{kl}[Q(\mu)||P(\mu|\tilde{s})]}_{\text{relative entropy}} - \underbrace{\ln P(\tilde{s})}_{\text{log of evidence}} \quad [14]$$

$$= \underbrace{D_{kl}[Q(\mu)||P(\mu)]}_{\text{complexity}} - \underbrace{E_Q[\ln P(\tilde{s}|\mu)]}_{\text{accuracy}} \quad [15]$$

Relative entropy cannot be less than zero, so that free energy is minimised when, in the Bayesian model, the approximate posterior becomes the true posterior and the free energy becomes the negative log of evidence for the internal sudo generative model (69). Through minimisation of free energy, the difference between the μ and \tilde{s} is minimised, as \tilde{x} is a component of \tilde{s} this results in environmental learning.

1137 The minimisation of relative entropy can be seen in figure 9
1138 as the state frequency variance of both the paddle and ball
1139 become closer as performance increases. Log of evidence is re-
1140 presentative of the hydrogel memory, as information is collected
1141 the memory becomes saturated leading to less impact on the
1142 overall system. Via the definition of equation 14, minimising
1143 free energy is equivalent to maximising model evidence, and
1144 equivalent to minimising the complexity of accurate expla-
1145 nations for sensations (68), this is shown in the equivalent
1146 equation 15, where E_Q is the expected outcome respect to Q .
1147 Minimisation of free energy equates to minimisation of com-
1148 plexity, observed in the simplification of distributions in figure
1149 7, and maximisation of accuracy, observed in the improved
1150 performance of figure 6.

1151 The experimental results, presented in figure 10, also show
1152 that when an incorrect representation of the virtual environ-
1153 ment is provided to the hydrogel the performance is severely
1154 inhibited. This behaviour can be explained through the mem-
1155 ory mechanic. The hydrogel seeks to represent the information
1156 given, by redistributing the ions so as to reduce free energy.
1157 If the information provided to the gel does not accurately
1158 represent the environment in which the gel's outputs are being
1159 applied, the ions will redistribute to represent this incorrect
1160 information. Thus the gel's sudo internal model of the envi-
1161 ronment will be wrong, and so the gel's outputs will be not
1162 perform the task accurately.

1163 Conclusion

1164 This study applied the theories of learning in BNNs through
1165 FEP to ionic EAP hydrogels, a medium whose behaviour
1166 is also driven by the theories of FEP. Firstly, by measuring
1167 ion concentrations through conductivity of the EAP hydrogel
1168 after periods of stimulation, the memory mechanics of the
1169 ionic EAP hydrogel were highlighted to provide a basis for
1170 further experimentation.

1171 Secondly, utilizing techniques in reservoir computing, the
1172 ionic EAP hydrogel is embodied in the simulated game world
1173 of Pong, through the use of a custom MEA. The game environ-
1174 ment was encoded into stimulations provided to the hydrogel
1175 and recorded ion concentrations were used as motor commands
1176 within the game world. Through analysis of the behaviour
1177 of the EAP hydrogel within this simulated environment, im-
1178 proved performance was observed through the course of the
1179 game. Additionally, through analysis of how performance
1180 improvement manifested through behavioural changes, the
1181 mechanisms of the EAP hydrogels were linked back to FEP.

1182 By investigating the chain of events leading from ion migra-
1183 tion to game performance, the way the theories of FEP present
1184 as learning in the case of ionic EAP hydrogels were clarified.
1185 This was achieved by recording the various layers of abstraction
1186 from electric field stimulation to rally length, analysing
1187 the connections between these layers. The Pong game acts as
1188 a problem to be solved, when this problem is provided to the
1189 hydrogel as an electric field it causes an excess of free energy.
1190 Through FEP the hydrogel created a sudo internal model of
1191 the environment via ion migration, effectively learning to play
1192 the game. The ions within the hydrogel seek to minimise this
1193 free energy and redistribute, the redistribution of ions then
1194 becomes the solution to the problem. The way in which the
1195 solution is interpreted allows it to be applied to the problem,
1196 in this case through conversation of localised ion concentration

1197 to paddle motion. As ions migrate to represent the game
1198 environment, presented as a series of changing electric fields,
1199 the memory in the hydrogel is filled resulting in an increase in
1200 current draw. This then causes a change in paddle motion to
1201 better intercept the ball, which in turn improves the hit/miss
1202 ratio in the regions, which in turn increases the rally lengths.
1203

1204 Lastly, a set of baseline experiments were also performed.
1205 The goal of these experiments was to show that for the EAP
1206 hydrogel to learn it must be presented with information accu-
1207 rate to the environment in which it is embodied. By altering
1208 the mapping between regions and stimulation locations the
1209 'vision' of the hydrogel was impaired. This impairment severely
1210 impacted the performance of the hydrogel within the game.
1211 The severe reduction in performance supports the claim that
1212 for the hydrogel, or any FEP based medium, to learn it must
1213 be supplied with information that accurately represents the
1214 problem being learned.

1215 **Future Work.** The results of this study demonstrating that a
1216 form of emergent learning is possible in mediums other than
1217 BNNs, when the medium in question is governed by the same
1218 underlying principles of free energy. The application demon-
1219 strated, however, does not achieve the same resolution of
1220 ability that can be achieved with BNNs. As future work higher
1221 resolution MEAs can be tested to investigate the limits of this
1222 form of reservoir computing, similarly additional applications
1223 can be attempted to investigate how the learning behavior
1224 is influenced by task. FEP is not only found in ionic EAPs
1225 and BNNs but many other forms of active matter and natural
1226 systems, through exploration of these alternatives a medium
1227 that is more capable, and just as accessible, could be found.
1228 All these avenues of research, however, fall outside the goals
1229 of this study but describe an interesting path for future work
1230 and application of this technology.

1231 **ACKNOWLEDGMENTS.** This project is fully funded by Process
1232 Vision Ltd (<https://www.processvision.com/>).

- 1 HJ Ruskin, R Walshe, Emergent computing-introduction to the special theme. *ERCIM News* **64**, 24–25 (2006).
- 2 S Bondar, JC Hsu, A Pfouga, J Stjepandić, Agile digitale transformation of enterprise architecture models in engineering collaboration. *Procedia Manuf.* **11**, 1343–1350 (2017).
- 3 C Fernando, S Sojakka, Pattern recognition in a bucket in *European conference on artificial life*. (Springer), pp. 588–597 (2003).
- 4 A Tero, et al., Rules for biologically inspired adaptive network design. *Science* **327**, 439–442 (2010).
- 5 M Dueñas-Díez, J Pérez-Mercader, How chemistry computes: language recognition by non-biochemical chemical automata. from finite automata to turing machines. *Iscience* **19**, 514–526 (2019).
- 6 A Wang, et al., Configurable nor gate arrays from belousov-zhabotinsky micro-droplets. *The Eur. Phys. J. Special Top.* **225**, 211–227 (2016).
- 7 J Misra, I Saha, Artificial neural networks in hardware: A survey of two decades of progress. *Neurocomputing* **74**, 239–255 (2010).
- 8 J Végh, How amadhl's law limits the performance of large artificial neural networks. *Brain informatics* **6**, 1–11 (2019).
- 9 K Mainzer, From embodied mind to embodied robotics: Humanities and system theoretical aspects. *J. Physiol.* **103**, 296–304 (2009).
- 10 A Cangelosi, J Bongard, MH Fischer, S Nolfi, Embodied intelligence in *Springer Handbook of Computational Intelligence*. (Springer), pp. 697–714 (2015).
- 11 VC Müller, M Hoffmann, What Is Morphological Computation? On How the Body Contributes to Cognition and Control. *Artif. Life* **23**, 1–24 (2017).
- 12 L Shapiro, S Spaulding, Embodied Cognition in *The Stanford Encyclopedia of Philosophy*, ed. EN Zalta. (Metaphysics Research Lab, Stanford University), Winter 2021 edition, (2021).
- 13 G Tanaka, et al., Recent advances in physical reservoir computing: A review. *Neural Networks* **115**, 100–123 (2019).
- 14 B Schrauwen, D Verstraeten, J Van Campenhout, An overview of reservoir computing: theory, applications and implementations in *Proceedings of the 15th european symposium on artificial neural networks*. p. 471–482 2007. pp. 471–482 (2007).
- 15 K Heiney, OH Ramstad, I Sandvig, A Sandvig, S Nicelle, Assessment and manipulation of the computational capacity of in vitro neuronal networks through criticality in neuronal avalanches in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. (IEEE), pp. 247–254 (2019).

16. P Aaser, et al., Towards making a cyborg: A closed-loop reservoir-neuro system. *ECAL 2017, Fourteenth Eur. Conf. on Artif. Life* **14**, 430–437 (2017).
17. P Mateos-Aparicio, A Rodríguez-Moreno, The impact of studying brain plasticity. *Front. cellular neuroscience* **13**, 66 (2019).
18. ZC Chao, DJ Bakkum, SM Potter, Shaping embodied neural networks for adaptive goal-directed behavior. *PLoS computational biology* **4**, e1000042 (2008).
19. DJ Bakkum, ZC Chao, SM Potter, Spatio-temporal electrical stimuli shape behavior of an embodied cortical network in a goal-directed learning task. *J. neural engineering* **5**, 310 (2008).
20. TB DeMarse, KP Dockendorf, Adaptive flight control with living neuronal networks on microelectrode arrays in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.* (IEEE), Vol. 3, pp. 1548–1551 (2005).
21. BJ Kagan, et al., In vitro neurons learn and exhibit sentience when embodied in a simulated game-world (2021).
22. K Friston, The free-energy principle: a unified brain theory? *Nat. reviews neuroscience* **11**, 127–138 (2010).
23. D Demekas, T Parr, KJ Friston, An investigation of the free energy principle for emotion recognition. *Front. Comput. Neurosci.* **14**, 30 (2020).
24. K Friston, The free-energy principle: a rough guide to the brain? *Trends cognitive sciences* **13**, 293–301 (2009).
25. T Parr, KJ Friston, Generalised free energy and active inference. *Biol. cybernetics* **113**, 495–513 (2019).
26. K Friston, The history of the future of the bayesian brain. *NeuroImage* **62**, 1230–1233 (2012).
27. P Schwartenbeck, et al., Evidence for surprise minimization over value maximization in choice behavior. *Sci. reports* **5**, 1–14 (2015).
28. G Nicolis, I Prigogine, *Self-organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations.* (John Wiley, New York), (1977).
29. J Bruineberg, E Rietveld, Self-organization, free energy minimization, and optimal grip on a field of affordances. *Front. human neuroscience* **8**, 599 (2014).
30. JS Kelso, *Dynamic patterns: The self-organization of brain and behavior.* (MIT press), (1995).
31. V Pasquale, P Massobrio, L Bologna, M Chiappalone, S Martinoia, Self-organization and neuronal avalanches in networks of dissociated cortical neurons. *Neuroscience* **153**, 1354–1369 (2008).
32. K Friston, et al., The anatomy of choice: active inference and agency. *Front. human neuroscience* **7**, 598 (2013).
33. RM Neal, *Bayesian learning for neural networks.* (Springer Science & Business Media) Vol. 118, (2012).
34. DC Knill, A Pouget, The bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS Neurosci.* **27**, 712–719 (2004).
35. T Isomura, K Kotani, Y Jimbo, Cultured cortical neurons can perform blind source separation according to the free-energy principle. *PLoS computational biology* **11**, e1004643 (2015).
36. T Isomura, K Friston, In vitro neural networks minimise variational free energy. *Sci. reports* **8**, 1–14 (2018).
37. T Parr, KJ Friston, The discrete and continuous brain: from decisions to movement—and back again. *Neural computation* **30**, 2319–2347 (2018).
38. J Gordon, S Amiri, General overview of neuronal cell culture. *neuronal cell culture* **2311**, 1–8 (2021).
39. M Marchetti, et al., Hydrodynamics of soft active matter. *REVIEWS OF MODERN PHYSICS Rev Mod Phys* **85**, 1143 (2013).
40. M Otake, Y Kagami, M Inaba, H Inoue, Motion design of a starfish-shaped gel robot made of electro-active polymer gel. *Robotics Auton. Syst.* **40**, 185–191 (2002).
41. K Nakajima, T Li, H Hauser, R Pfeifer, Exploiting short-term memory in soft body dynamics as a computational resource. *J. The Royal Soc. Interface* **11**, 20140437 (2014).
42. R Kürsten, V Sushkov, T Ihle, Giant kovacs-like memory effect for active particles. *Phys Rev Lett* **119**, 188001 (2017).
43. T Tanaka, I Nishio, ST Sun, S Ueno-Nishio, Collapse of gels in an electric field. *Science* **218**, 467–469 (1982).
44. P Atkins, J De Paula, Atkins' physical chemistry, volume 8. ed (2006).
45. VV Yashin, S Suzuki, R Yoshida, AC Balazs, Controlling the dynamic behavior of heterogeneous self-oscillating gels. *J. Mater. Chem.* **22**, 13625–13636 (2012).
46. O Kuksenok, VV Yashin, AC Balazs, Mechanically induced chemical oscillations and motion in responsive gels. *Soft Matter* **3**, 1138–1144 (2007).
47. D De Tommasi, G Puglisi, G Zurlo, Hysteresis in electroactive polymers. *Eur. J. Mech.* **48**, 16–22 (2014).
48. M Bassil, J Davenas, ME Tahchi, Electrochemical properties and actuation mechanisms of polyacrylamide hydrogel for artificial muscle application. *Sensors Actuators B* **134**, 496–501 (2008).
49. JR Gray, Conductivity analyzers and their application. *Environ. instrumentation analysis handbook* **1**, 491–510 (2004).
50. P Atkins, J De Paula, *Physical chemistry for the life sciences.* (Oxford University Press, USA), (2011).
51. P Atkins, PW Atkins, J de Paula, *Atkins' physical chemistry.* (Oxford university press), (2014).
52. T Tanaka, et al., Phase transitions in ionic gels. *Phys. Rev. Lett.* **45**, 1636 (1980).
53. M Oyen, Mechanical characterisation of hydrogel materials. *Int. Mater. Rev.* **59**, 44–59 (2014).
54. R Bourchouladze, *Memories are made of this: How memory works in humans and animals.* (Columbia University Press), (2002).
55. G Schiavone, et al., Guidelines to study and develop soft electrode systems for neural stimulation. *Neuron* **108**, 238–258 (2020).
56. K Jia, X Li, Y Wang, Electrochemical breakdown in hydrogel ionotronic devices. *Soft Matter* **17**, 834–839 (2021).
57. Mathworks, Polyfit (2014) <https://www.mathworks.com/help/matlab/ref/polyfit.html#bu6sxq-1-S>. Accessed 01/10/2021.
58. T Viering, M Loog, The shape of learning curves: a review. *arXiv preprint arXiv:* **2103**, 10948 (2021).
59. N Leibowitz, B Baum, G Enden, A Karniel, The exponential learning equation as a function of successful trials results in sigmoid performance. *J. Math. Psychol.* **54**, 338–340 (2010).
60. S Siegal, *Nonparametric statistics for the behavioral sciences.* (McGraw-Hill), p. 121 (1956).
61. JH Zar, *Biostatistical Analysis.* (New Jersey: Prentice Hall International, INC.), p. 147 (1998).
62. H Lehmann, Erich; D'Abra, *Nonparametrics: Statistical Methods Based on Ranks.* (Holden-Day), p. 20 (1975).
63. Y Mao, M Cates, H Lekkerkerker, Depletion force in colloidal systems. *Phys. A: Stat. Mech. its Appl.* **222**, 10–24 (1995).
64. HN Lekkerkerker, R Tuinier, Depletion interaction in *Colloids and the depletion interaction.* (Springer), pp. 57–108 (2011).
65. DV Schroeder, *An introduction to thermal physic.* (San Francisco, CA: Addison Wesley) Vol. 57, (2000).
66. N Ebrahimi, E Maasoumi, ES Soofi, Ordering univariate distributions by entropy and variance. *J. Econom.* **90**, 317–336 (1999).
67. K Friston, G Buzsáki, The functional anatomy of time: what and when in the brain. *Trends cognitive sciences* **20**, 500–511 (2016).
68. K Friston, T FitzGerald, F Rigoli, P Schwartenbeck, G Pezzulo, Active inference: a process theory. *Neural computation* **29**, 1–49 (2017).
69. MJ Beal, *Variational algorithms for approximate Bayesian inference.* (University of London, University College London (United Kingdom)), (2003).