A Precarious Sensorimotor Sequence Reiterator for Modelling Enactive Habits

Felix M. G. Woolford[1], Matthew D. Egbert[1]

[1] School of Computer Science, University of Auckland, Auckland, New Zealand
fmg.woolford@gmail.com

Abstract

In the enactivist framework, *habits* are precarious, self-sustaining, and self-individuating sensorimotor structures: they are a first approximation of *autonomous* sensorimotor entities. We present a prototype computational model which demonstrates a relatively simple mechanism for facilitating the emergence of such habits in robots. At its core this model is a system which retains a history of sensorimotor sequences which are compared and re-enacted by the controller as a function of the recent sensorimotor activity of the controlled robot. To demonstrate an application of this model concretely, we also present a minimal cognition task loosely inspired by the role of sensorimotor contingencies in human colour perception. This task requires that a robot maintains a set of particular sensorimotor coordinations which allow it to respond to different objects appropriately, influenced by an evolved behavioural history.

Introduction

Recent work developing enactive theories of sensorimotor autonomy, sensorimotor agency and sensorimotor life (Barandiaran, 2017; Barandiaran et al., 2009; Di Paolo et al., 2017) has provided a new framework for understanding cognition in both natural and artificial systems. As part of these developments, a renewed emphasis has been placed on a rich notion of *habit* (Barandiaran and Di Paolo, 2014; Egbert and Barandiaran, 2014). The enactive notion of a habit is one of precarious, self-maintaining sensorimotor structures which form the basic units of a sensorimotor individual, analogous to the cells of biological lifeforms. Computational modelling has an important role to play in investigating and testing these theories. A prominent computational model for exploring habit-based behaviour which heavily influences this paper is the IDSM (Egbert and Barandiaran, 2014; Egbert and Cañamero, 2014; Egbert, 2018). Recent work utilizing the IDSM include using it to illustrate behavioural scaffolding through motor babble (Zarco and Egbert, 2019), and a variation which abstracts away from explicit sensorimotor dimensions to investigate decision-making processes (Batta and Stephens, 2019). Additionally a variation of the IDSM has been developed which allowed the enumeration some basic behaviours which are achievable with such a controller (Woolford and Egbert, 2019).

In order to broaden the scope of possible investigations into habit and sensorimotor agency, we are developing a new model which offers an alternative approach to modelling habits. We present a prototype model called a *sensorimotor sequence reiterator* (SSR) which is intended as a simple platform for investigating the self-generation and maintenance of a network of habitual behaviours. Whereas the IDSM is based on historically-determined plastic mappings of sensorimotor-state to motor-action, this model offers a different approach: generating activity through the construction and traversal a network of historical sensorimotor sequences. The SSR is has unique advantages which warrant it to be investigated in its own right. In particular, the SSR is designed to be simpler in conceptual and analytical terms than the IDSM, which has often proven to be a challenging model to incorporate into new research programmes. Additionally, due to its operational mechanism, the SSR has a different domain of practical applications. Specifically it may be more suited to investigating networks of multiple habits because moment-to-moment motor activation has a stronger history-dependence than IDSM motor activation. In this paper we discuss the functional mechanism of the model and present a simple investigation with it in order to demonstrate its potential as an experimental tool. In this task the robot is required to exploit a handful of contingent relationships between its motor and sensor activity over time in order to distinguish between coloured objects and respond appropriately. To bias the robot towards achieving this goal, we use a method of transferring heritable behaviour through experience, as opposed to the parametric tuning of classical minimal cognition evolutionary robotics experiments.

Model

The SSR is a robot controller which functions by comparing recent sensorimotor activity to past activity and attempting to repeat the most similar historical behaviour. The core element of the SSR is what we will refer to as its *historical sequence*, which is its representation of the recent history of its controlled robot's discrete sensorimotor-states (SM-states) as a sequence over time. A robot's SM-state is represented as the pair of a particular motor action and the sensory state which results from enacting that action in the environment, a system loosely motivated by Maye and Engel (2011). As a minimal illustration of this, let us consider a robot which moves around a 2-dimensional white and black checkerboard-like environment in discrete steps (U, L, D, R), and senses the colour of the square on which it sits (w, b). If the robot were to start on a black square and take four moves in the order up, left, down, and left again, its sequence of SM-states from the first action would be represented as $wU \rightarrow bL \rightarrow wD \rightarrow bL$, as illustrated in Figure 1.

A SSR builds its historical sequence as it operates, uses it to determine future actions, and modulates it as a function of those actions. Each individual SM-state in the historical sequence is represented by a node, which describes the SM-state in terms of an instantaneous sensory state S and a motor action M, and also represents the motor action taken from that SM-state, O, and a plastic weighting W which is a function of both the node's age and the relationship between itself and previous nodes. Thus an individual node may be expressed as a tuple $N = \langle S, M, O, W \rangle$, and can be thought of as an instance of a historical motor-state transition with an intermediary sensory state. Each node is linked to from another node representing the previous state and itself links to a node representing the next SM-state. Nodes survive for a finite amount of steps, H, and are then destroyed, so the historical sequence is essentially a shifting window across the robot's H most recent states.

We now explain how the historical sequence is generated. Nodes are constructed as the SSR's robot interacts with its environment. Each iteration a new node is generated in the following process: First, the SSR selects a motor action for the robot to take in the current time step. If there is already an active node this action is set as the node's outward action. The robot then enacts the selected action in the environment, and its sensory state is updated. The pairing of the motor action and resulting sensory state is treated as the robot's instantaneous SM-state. The current active node becomes inactive and is linked to a newly generated active node which represents the new SM-state. This is true even if the new SM-state is identical to the previous. The new node's outward action is left undefined until it is determined in the next iteration, as in the fourth node in Figure 1.

The SSR chooses which action to take in a manner which causes previous sequences of behaviour to be reiterated when they resonate with the regularities of the robot's engagement with its environment. Motor activation is a function of the similarity of the current node to past nodes and the weight of the most similar past nodes. The SSR finds the node which maximizes these elements and selects the action that was taken from that past node to enact again in the

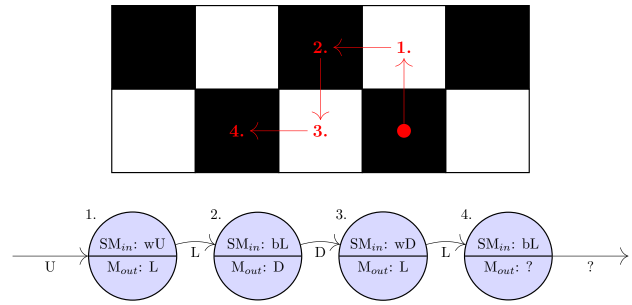

Figure 1: A sequence of actions in the environment and the corresponding nodes in the SSR's historical sequence.

current iteration, with a degree of randomisation. Of these elements, the similarity metric is the first and most significant. Similarity between present and past nodes is evaluated in terms of the sequence of states leading up to those nodes. Individual nodes may share the same SM-state with other nodes, and distinct subsequences of nodes which successively share SM-states are said to be more similar the longer they are. During motor activation, the SSR evaluates those past nodes which have the same state as the active node in an attempt to find a *matching node* which will determine the selected action. Potential matching nodes are ranked based on the similarity of the sequence of their preceding nodes to the sequence of nodes that precede the active node.

The process of evaluating the similarity of matching nodes is as follows: For a given active node N_a with a particular SM-state, there are any number of nodes with the same SM-state in the historical sequence. Let us assume for the moment that there is at least one potential matching node and call one of them N_b. To compare the similarity of N_a to N_b, we must examine the states of the preceding nodes of each. Nodes are compared in reverse order, and we describe as N_{a-1} the node preceding N_a and so on. N_{a-1} is compared to N_{b-1}, N_{a-2} is compared to N_{b-2}, et cetera. Similarity is measured in terms of the length of the sequence of each node's predecessors with identical SM-states, so when the SM-state of N_{a-x} is dissimilar to that of N_{b-x}, and these are the first dissimilar nodes in the reverse sequence, N_a and N_b are said to have a similarity of $x-1$. More formally, two nodes N_a and N_b are said to have a similarity of η if:

$$\begin{aligned} &N_{a-i}^{SM} = N_{b-i}^{SM},\ i \in \mathbb{Z} \,|\, 0 \leq i \leq \eta, \\ &N_{a-(\eta+1)}^{SM} \neq N_{b-(\eta+1)}^{SM} \end{aligned} \tag{1}$$

Where the superscript SM refers to taking only the S and M elements of the tuple N_x. The past node with the greatest similarity is selected as the best match. This mechanism often yields ties between matching nodes with different outward actions, and in order break ties we implement a weighting mechanism. A node's weight is initially set as an increment of the weight of its matching node during creation, and

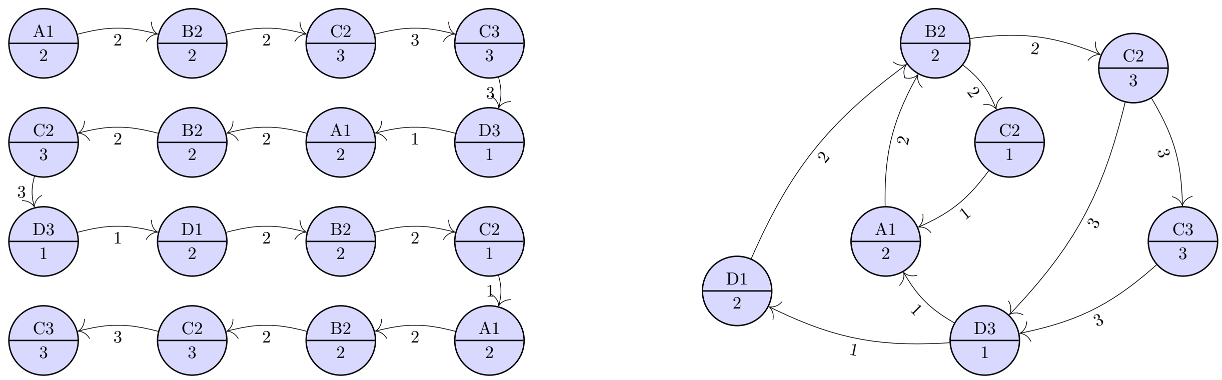

Figure 2: Example of a network generated by treating nodes as equivalent to their matches, using abstract sensor states A, B, C, D and motor states $1, 2, 3$. The left part of the figure illustrates 16 nodes as they appear in the SSR's historical sequence, the network on the right illustrates how subsequences lead back to their own re-enactment, creating cycles.

decreases with age. This weight is formally defined as

$$W = (W_m + B) \times \left(1 - \frac{1}{e^{\frac{10(H-\tau)}{H}}}\right) \tag{2}$$

where W_m is the weight of the node which was its match at the time of its generation; B is a parameter which defines the amount that the new node's weight is incremented over its matching node's weight; τ is the time since the creation of the node in discrete steps; and H is the aforementioned maximum age of a node before it is destroyed. This weight is intended to capture the notion of reinforcement of behaviour through repetition. Weights are propagated when nodes pass their action on to future nodes, and thus nodes which represent actions with a long history have heavier weights. When multiple matching nodes have equal similarity, that one which has the greatest weight is selected as the best node. Since the decay function is continuous, further ties are unlikely, but any final tie-breaking favours older candidate as the best matching node.

The final factor influencing motor activation is the random exploration element. This is a small chance that the controller will ignore the action of the selected node and instead perform a random action, uniformly selected from the set of all possible actions. The probability of this occurring depends on the similarity of the best matching node to the active node and is given as

$$P(r) = \max\left(0.5\left(1 - \frac{\eta}{D}\right), 0.005\right) \tag{3}$$

where η is the similarity of the best matching node and D is a parameter indicating the similarity at which the minimum probability of selecting a random action is achieved. Note that if there is no matching node in the historical sequence then $P(r) = 1$. When a random action is selected, the previously found best matching node is ignored and a different node is found to be considered the best match for the purposes of assigning a weight to the active node. This is achieved by finding all historic nodes which have the same SM-state as the active node and which also have the same outward action as that which was selected randomly. Of these nodes, that which is the most similar and has the strongest weight lends its weight to the active node, even if it is less similar and/or weaker than the original best match. Note that it is possible that the action that is selected randomly will be the same as that of the original matching node. In this case that same node will be reselected as the best match.

The typical behaviour of a SSR changes as it matures. This is due to the circular dependency in the way that the historical sequence is generated as the SSR selects actions for its robot at the same time as motor activation is determined by the historical sequence. In its simplest form, a SSR begins with no behavioural history to draw from. Therefore the initial phase of its behaviour is mostly random, with active nodes having either no matching nodes or relatively dissimilar nodes in the historical sequence. This is an important quality because it means that the SSR has an inherent tendency to more freely explore the potentials of its sensorimotor environment in its early stages, akin to motor babbling in infants (see also Zarco and Egbert, 2019). As the historical sequence grows larger, the SSR's behaviour becomes gradually more deterministic as repeatable patterns of activity form.

Because nodes in the historical sequence have a limited lifespan, behaviours which have not been enacted for too long are lost. When a sequence of states is reiterated, the behaviour itself survives in the historical sequence even after the original nodes representing the behaviour are forgotten. As a simple example, the sequence $Uw \rightarrow Lb \rightarrow Dw \rightarrow Rb \rightarrow Uw \rightarrow Lb \rightarrow Dw \rightarrow Rb$ may be treated as two iterations of the smaller sequence $Uw \rightarrow Lb \rightarrow Dw \rightarrow Rb$. This equivalence is captured in the SSR's weighting mechanism, which is intended as a reflection of the frequency and recency of an enactment of a particular behaviour. Let us number these states from 0 to 7: Assuming that these are the first states in the history sequence, if the weights of the nodes representing states 0 to 3 have a value of 1, then the weights of nodes representing 4 to 7 will have weight $1 + b$, reflecting the fact that those latter states are a reenactment of the former. Through this mechanism, behaviours which are repeated more frequently are more likely to be repeated in the future because their weights are compounded. Furthermore, by treating those nodes which represent reenactments of previous states as equivalent to those previous states' nodes, we may recast the linear historical sequence as a network with cyclical subsequences, as in Figure 2. The significance of this will be addressed in the final section.

Investigation and results

To briefly illustrate the capacities of the SSR and to provide a case in which to analyse it behaviour, we now present an investigation in which a robot controlled by a SSR reliably performs a minimal object discrimination task. This experiment is loosely inspired by investigations into the relationship between eye saccades and colour perception (Bompas and O'Regan, 2006; O'Regan and Noë, 2001). In a time-discrete simulation, the robot is presented with an environment filled with a mix of three varieties of objects coloured either red, green, and blue. The robot may move around the environment in a limited fashion by adjusting its bearing as it moves forwards steadily, and when it moves over an object that object is 'eaten' and disappears. The robot is tasked with eating the blue objects while avoiding the red and green objects. The task is made less trivial due to the limited sensory information available to the robot: There are only four possible sensory states available to the robot, states L, F, R, and N, which respectively correspond to the robot detecting an object in its left sensor, front sensor, or right sensor, or detecting no object at all. The robot has a limited ability to distinguish between object colour, based on the way its sensors are triggered as it moves. Specifically, the robot may only detect red and blue objects when it is turning right, and may only detect green and blue objects when it is turning left. It detects all three varieties of objects when it is not turning at all. The sensors themselves do not provide information regarding the object's colour, for instance if the robot detects an object in its forward sensor while turning left, its sensory state will be F regardless of whether that object is blue or green. This sensory mechanism is illustrated further in Figure 3.

The effect of this mechanism is that no instantaneous state information is sufficient to indicate the exact colour of a detected object, and in order to identify an object and hence respond to its stimulus 'correctly' the robot must employ an active exploratory strategy to integrate historical sensory states. We will see shortly that the SSR is capable of developing such a strategy, after we briefly describe the precise details of the experiment.

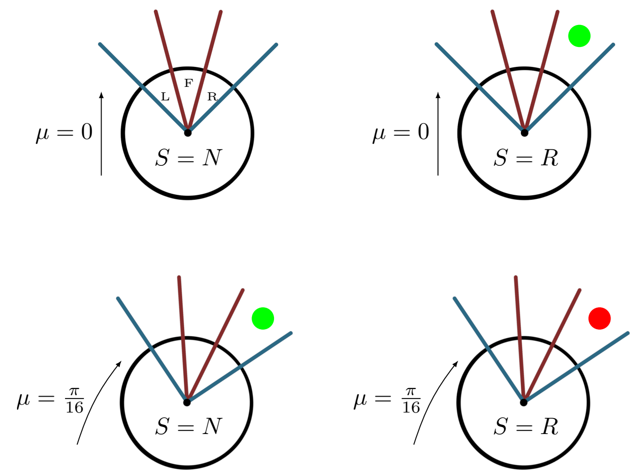

Figure 3: Examples of the relationship between actions, perceptions, and the environment for the investigated robot. Top left: No object is triggering any sensor. Top right: A green object triggers the robot's right sensor while it is not turning either direction. Bottom left: A green object is in the robot's right sensor's arc but the robot is turning right so the sensor is not triggered. Bottom right: The same scenario but with a red object, which does trigger the sensor when the robot is turning right.

The robot's body is a circle with a radius of 3 units, and as mentioned it has 3 sensors which each detect objects within a fixed sector and a range of 28 units. The front sensor detects any object in the sector within $\frac{\pm\pi}{12}$ radians of the front point; The left sensor detects objects between $\frac{-\pi}{4}$ and $\frac{-\pi}{12}$ radians away from the front point, and the right sensor detects those objects within the sector between $\frac{\pi}{4}$ and $\frac{\pi}{12}$ away from the front point. The robot has a single motor under its control, which rotates the robot by a fixed amount in a single time step. There are 7 possible states of this motor: 3 corresponding to light, moderate, and hard left rotations, 3 corresponding to equivalent rightward rotations, and 1 which corresponds to no rotation. Additionally to rotation, the robot is driven forward in the direction of its bearing at a steady rate of 1 unit per time step, but this is not under the control of the SSR. Each time step proceeds in the following order: First the robot rotates according to its motor state, it then moves forward one unit, it then eats any object within its bodily radius, and finally it updates its sensory state. After this sequence the SSR updates its historical sequence and selects a rotation action for the robot to take in the next step.

The arena is a two dimensional field of size 400 by 400 units with toroidal boundary conditions, and its only features are the 81 objects, divided evenly among the three colours, positioned along a regular grid 40 units apart. This regular positioning simplifies the model by ensuring that the robot

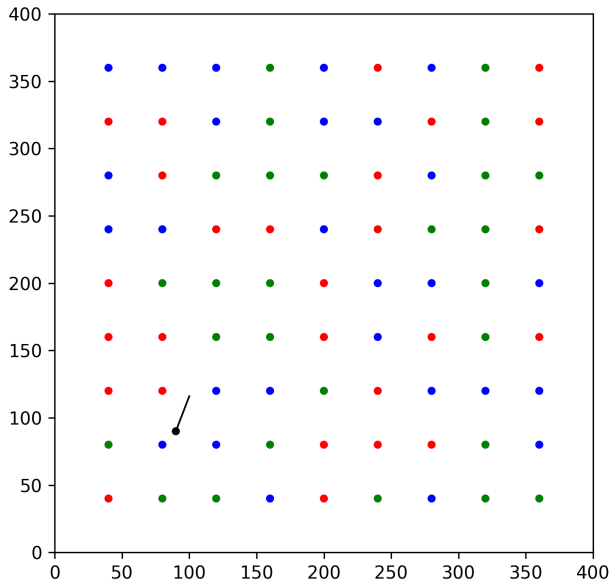

Figure 4: The robot (black object) in an example initial environment. Objects disappear when they are eaten and are not replaced.

can never detect more than one object at any time. The arrangement of the colours is shuffled randomly at the start of each experimental trial so as to prevent the success of rote trajectories through the environment. An example environment is illustrated in Figure 4. Finally, the SSR's parameters as per Equations 2 and 3 are set by hand to $H = 4000$, $B = 1$, and $D = 6$.

We now explain how a particular SSR may be used to control a robot to reliably perform the desired task. Despite the similarity to a typical evolutionary robotics minimal cognition experiment, standard evolutionary techniques are not suitable for selecting and optimising a particular configuration to a desired task. Unlike a neural network, for instance, the SSR has intrinsic behavioural dynamics which cannot be readily manipulated through parametric optimisation. Instead, we use an initial sequence of states to seed the behaviour of the SSR, as opposed to the default of having the SSR start with an empty history. A candidate solution therefore is a sequence of states which can be inserted into the start of a SSR's historical sequence over many different trials, causing it to reliably produce behaviour in which the robot eats blue objects and avoids the others. Since the only inherited behaviour comes in the form of an historical 'experience', this demonstrates a relatively unexplored method of influencing robot behaviour in a manner which could be considered analogous to instinctive behaviour or that acquired through parental or social scaffolding.

To find solutions, we used a simple genetic algorithm. For the first generation of 700 individuals the SSRs operated as normal to generate their historical sequences. The first 750 nodes in these sequences then formed a genome which was passed along to descendant individuals. The fitness of the solution was measured by a per-trial fitness function of

$$F = b - r - g \tag{4}$$

where b, r, and g are the number of blue, red, and green objects eaten over the course of a trial of 10000 time steps. This fitness value was averaged across 10 trials, with randomised start positions and arrangements of objects, to determine the final fitness of the individual. Despite the naivety of the implementation, this search yielded candidates which were sufficient for our illustrative purposes here.

After using the evolutionary algorithm to select a genome of initial actions which yielded a good average fitness, we then ran 10 further simulations of 10000 steps with the best candidate genome, and selected a random simulation for further analysis. This particular simulation yielded a fitness of $F = 22$ according to Equation 4, with $b = 22$ and $r = g = 0$. Note that this was slightly atypical as in most cases the robot would eat a few red and/or green objects.

We will uses results from this investigation to illustrate the internal behaviour of the SSR, highlighting: how certain behaviours repeatedly *resonate* with (i.e. are drawn out by- and suitable for-) particular environmental regularities; how the structures which form enable the robot to perform the task; and how certain behaviours sustain themselves through re-enactment and how others fail to do so.

This section is intended to be illustrative, rather than an exhaustive analysis of all of the possible behaviours with the SSR in this simulation. Thus we will isolate a handful of useful examples to discuss here.

As discussed earlier, a linear sequence of nodes can be presented as a digraph such as the one in Figure 5, where each vertex represents a group of equivalent nodes, and an edge between any vertices u and v represent a transition from any node represented by u to any node represented by v. We are able to see that there are cycles throughout the graph, and furthermore in most cases these cycles each tend to show a high degree of context sensitivity, i.e. they are enacted in a single environmental condition such as when a particular-coloured object is within the robot's scanning range. These cycles form interesting structures - they emerge through the robot's exploration of its environment, they cause the robot to re-establish the conditions for their own enactment, and they are revisited when the regularities of the robot's sensorimotor engagement resonates with them. There are three distinctly blue structures in the graph, and we will focus on the structure detailed in Figure 6, which represents the behaviour most frequently enacted on order to eat a blue object. We will refer to this as the $\alpha-$structure. Over the 10000 steps of the simulation, the robot encountered a blue object 42 times. By 'encounter'

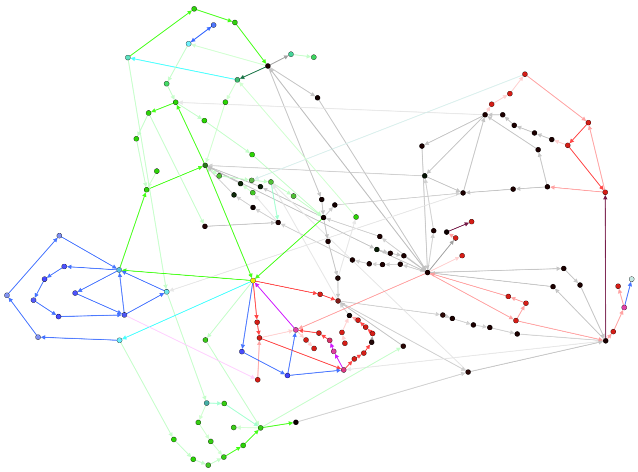

Figure 5: A network with vertices representing equivalent nodes for the presented simulation, as in Figure 2. Vertices which are visited fewer than 5 times have been excluded. RGB values indicate the objects which are encountered as edges are traversed and nodes visited - for instance a deep blue indicates that the robot is always encountering a blue object at a particular vertex, green indicates that always green objects are encountered, and intermediate shades indicate an mixture of encounters.

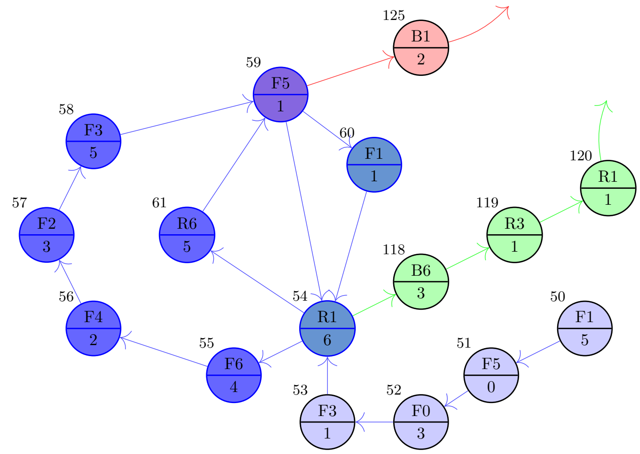

Figure 6: A detail from the network in Figure 5 illustrating the $\alpha-$structure (vertices 54 to 61) and other closely associated vertices. The vertices are coloured as in Figure 5. Vertices 50 to 53 are a subsequence which is not part of the cycles in the $\alpha-$structure but frequently lead into it (see Figure 7), and vertices 125, 118, 119, and 120 are parts of subsequences which lead away from the structure.

we refer to an unbroken sequence of steps in which there is an object within the robot's scanning range. 12 of these encounters were extremely short - only lasting 2 or 3 steps. Of the remaining 30 encounters, 22 ended when the robot eats the object. Figure 7 illustrates the robot's behaviour in those 22 encounters. The majority of these approaches are dominated by repeated enactments of behaviour described by the $\alpha-$structure.

In the network, the $\alpha-$structure consists of four intersecting cycles which consist of left, right, and null turns, but only F and R sensory states. Qualitatively the behaviour of the robot in this structure can be described as scanning left and right in a manner which distinguishes the colour of the encountered object, and moving forward towards blue objects if they are to the front, or bearing rightwards if the object is to the right. An interesting feature of the structure is the junction from vertex 54 leading to either vertex 55 or 61. The subsequence of $54 \rightarrow 55 \rightarrow 56 \rightarrow 57 \rightarrow 58 \rightarrow 59$ describes a gentle rightward turn when the object is detected to be approximately to the fore, whereas the subsequence of $54 \rightarrow 61 \rightarrow 59$ describes a stronger rightward turn when the SM-state indicates that the object is further to the right. The alternation of these two subsequences allows the robot to adjust its bearing appropriately as it approaches the object. The alternation of left and right turns in the structure allows the robot to distinguish the colour of the encountered object and take evasive action when the object is not blue. Vertices 54, 59, and 60 are occasionally visited when the robot encounters green or red objects, and are connected to other sequences which are enacted in those cases. As an example, vertex 54 is visited 18 times when the robot is encountering a green object, and 67 times when it is encountering a blue object. However of the three vertices which succeed 54, 55 and 61 are only visited when the robot is encountering a blue object and 118 is only visited when the robot is encountering a green object. Thus we can localise a sort of discriminating behaviour occurring around vertex 54, whereas other vertices in the $\alpha-$structure could be said to visited when the robot is adjusting its bearing to head closer to blue objects. It is important to note that this behaviour at vertex 54 is due to its relationship to preceding vertices, not simply the robot's instantaneous SM-state. At vertex 54, the robot's SM-state is $1R$, which indicates a hard leftward turn leading to a triggering of the right sensor, and from vertex 54 the robot a slight right turn. However at vertex 120 the robot has the same SM-state but takes a different action - a second hard left turn. This is an advantageous behaviour for the task at hand: Because vertex 120 is in a sequence which resonates when the robot encounters a green object, the subsequent action to take another leftward turn serves to make it more likely that the robot will avoid green objects. On the other hand since vertex 54 is visited in sequences which resonate when the robot encounters blue or green objects, the subsequent rightward turn serves to help distinguish blue and green objects and therefore lead to sequences which cause the robot to avoid or approach the object depending on the sensorimotor consequences of that action.

The final aspect of the SSR we will illustrate here is the

	$F=b-r-g$	b	r	g
before objects removed	18.6	22.1	2.11	1.47
after objects restored	-7.60	10.2	10.1	7.64
no initial history	-4.79	5.47	5.02	5.24

Table 1: Mean number of objects eaten over 200 trials. The third row provides a baseline comparison for robots without the evolved initial history.

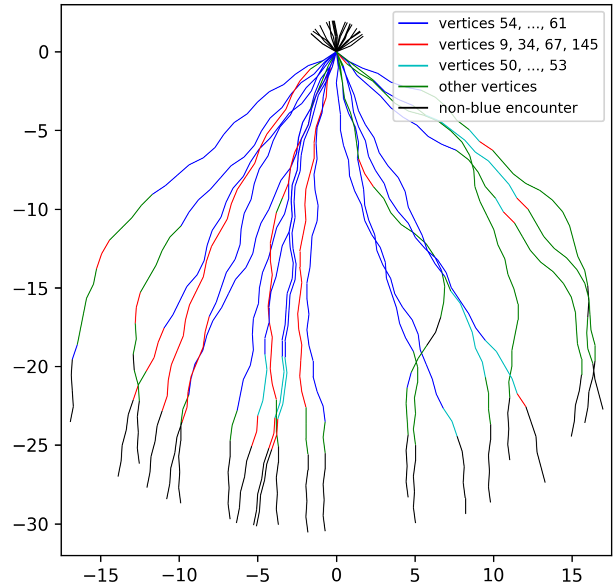

Figure 7: Robot's trajectories for the 30 steps leading up to each of the 22 blue object consumptions. Trajectories are transposed such that the eaten object is at position (0, 0) and the initial bearing is north. Coloured segments indicate the relationship between the SSR's active node and the vertices on the graph in Figures 5 and 6.

precariousness of the emergent behavioural structures. In order to illustrate the effects of this precariousness, we ran a short additional investigation to examine the effect of disrupting the regular sensorimotor contingencies of the robot's engagement with the environment. Using the same initial historical sequence as evolved in the main investigation, we ran 200 additional simulations of 30000 steps, divided into three phases: The first 10000 steps were identical to the previous simulations (Phase 1). At step 10000, all objects were removed and the robot continued to move around an empty environment (Phase 2). At step 20000, all objects (including those previously eaten) were restored to their original position, and the robot's fitness was reset to 0 (Phase 3). We averaged the fitnesses from Phase 1 and Phase 3 over the 200 simulations, and as expected find a significant decrease in fitness in Phase 3 compared to Phase 1. This is a result of the useful structures - which were initially found in the evolved genome and then sustained through regular engagement with the coloured objects - being lost when they cannot be reinforced in the middle phase of the simulation. iSince most of the structures formed in the first phase are not able to be reinforced during the second phase, the robot loses its ability to respond to the objects in the way we desire when they are restored in the third phase. The results of these trials are presented in Table 1.

Discussion

The enactive habit is an entity which is self-individuating, precarious, and self-reproducing. The SSR investigated here supported several such entities. At the lowest level, we see self-individuation with the way that subsequences which are repeated successively become increasingly likely to repeat themselves ever further. In the projection in Figure 5 these subsequences are visible as cycles. By their cyclical nature they are inherently self-reproducing because they create the conditions for their own re-enactment by repeating earlier SM-states. Some such cycles are more readily revisited than others, because they resonate with more frequently encountered environmental regularities. Those cycles which restore themselves reliably, and which are revisited frequently, form the core of the SSR's overall structure. While these are cases of minimal self-individuation, there are also instances - such as in the case of the $\alpha-$structure - where multiple cycles may be tightly associated with one another and frequently lead into each other, and with clear boundaries (e.g. vertices 59, 60, and 61 in Figure 6) which connect to other structures which are not as intertwined. These structure's are what we consider to be the locus of the *habits* in the SSR. In addition, there are important subsequences which do not constitute habits as they do not regularly cycle back on themselves, but are nevertheless significant to the overall behavioural structure. The sequence of vertices 50 to 53 are an example of this (illustrated in Figures 6 and 7). In all but one case this sequence is not repeated in succession, but in all cases it leads into the $\alpha-$structure. In effect it represents a transitional structure between habits.

The averages in Table 1 illustrate the precariousness of the habits. In that experiment, some structures are lost because the robot can no longer interact with the objects, whereas the structures which resonate when the robot sees nothing are reinforced. This has interesting consequences for how the behaviour of the system changes over time. In addition to the obvious consequences discussed in the previous section, we also see in Table 1 that the evolved robot eats more objects of all colours in the third phase than the unevolved ones eat. This is likely due to the fact that some structures are reinforced through the second phase, specifically those corresponding to behaviour when the robot encounters no objects. In the evolved robot, this behaviour tends to be somewhat more efficient for exploring and finding new objects, whereas the unevolved robots have a strong tendency to move in circles. For similar reasons, the evolved robot shows a slightly different response to green objects. This is due to the fact that certain subsequences of behaviour can be enacted when the robot is encountering red or green objects or no object at all, and hence can be reinforced even without encountering objects.

Our investigation also explores the relationship between habits and broader behavioural strategies. In addition to being an example habit, we have focussed our analysis on the $\alpha-$structure because it is the most clear example of a structure in the behavioural network enabling the robot to perform the task we have set it. The structure could potentially be understood as a sort of Piagetian sensorimotor scheme (Boom, 2010; Di Paolo et al., 2017). It represents a small repertoire of low level sensorimotor activity which allows the robot to perform the higher order action of 'eating a blue object' under a variety of similar but distinct conditions. Linked to other structures, such as those in which the robot bears away from non-blue objects, there is a clear progression from individual sensorimotor states up to a cohesive behavioural strategy.

In our investigation we are able to, in a sense, evolve an adaptive behaviour through generating an initial historical sequence. However the use of the genetic algorithm means that the adaptive element is extrinsic to the SSR mechanism itself, and requires imposed viability constraints which are not directly related to the formation of habits (see also Beer, 1997, p.265). An open problem with understanding habits in terms of sensorimotor agency is the question of how a habit itself can be adaptive. In order to move towards a model of sensorimotor agency, one approach would be to extend the model to one which has a sort of 'habitual metabolism' and which alters its structure to adaptively regulate this metabolism in the face of changing environmental regularities. The SSR in the minimal form presented here does not capture such a notion of adaptivity at all, but we believe there is potential for it to be developed in this direction and should be the next step in research involving this model. Additionally, the results in Table 1 suggest that there is potential to use this model to investigate the relationship between habit and behavioural development, in that exposing the robot to particular environments sequentially may influence the SSR's response to later environments. As a minimal example, if the robot has historically interacted with a green object before it interacts with a blue object, its response to the blue object will potentially be influenced in part by its prior interaction with the green object. Taken further, this may be the basis for an interesting way of manipulating the robot's behaviour in a more developmentally meaningful way than the typical optimisation approach employed here. nother step could be to explore developing a system of clustered SSRs, which are in competition for control of their robot and which possess a mechanism through which they modulate their own historical sequences in response to the current sensorimotor environment. Such a cluster could provide a step towards modelling an adaptive ecology of habits (eg. Di Paolo et al., 2017).

References

Barandiaran, X. E. (2017). Autonomy and Enactivism: Towards a Theory of Sensorimotor Autonomous Agency. *Topoi*, 36(3):409–430.

Barandiaran, X. E., Di Paolo, E., and Rohde, M. (2009). Defining Agency: Individuality, Normativity, Asymmetry, and Spatio-temporality in Action. *Adaptive Behavior*, 17(5):367–386.

Barandiaran, X. E. and Di Paolo, E. A. (2014). A genealogical map of the concept of habit. *Frontiers in Human Neuroscience*, 8.

Batta, E. and Stephens, C. (2019). Heuristics as decision-making habits of autonomous sensorimotor agents. *Artificial Life Conference Proceedings*, (31):72–78.

Beer, R. D. (1997). The dynamics of adaptive behavior: A research program. *Robotics and Autonomous Systems*, 20(2-4):257–289.

Bompas, A. and O'Regan, J. K. (2006). Evidence for a Role of Action in Colour Perception. *Perception*, 35(1):65–78.

Boom, J. (2010). Piaget on Equilibration. In Müller, U., Carpendale, J. I. M., and Smith, L., editors, *The Cambridge companion to Piaget*, pages 132–149. Cambridge University Press, Cambridge.

Di Paolo, E., Buhrmann, T., and Barandiaran, X. (2017). *Sensorimotor Life: An enactive proposal*. Oxford University Press.

Egbert, M. (2018). Investigations of an Adaptive and Autonomous Sensorimotor Individual. In *The 2018 Conference on Artificial Life*, pages 343–350, Tokyo, Japan. MIT Press.

Egbert, M. and Cañamero, L. (2014). Habit-Based Regulation of Essential Variables. *Artificial Life 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, pages 168–175.

Egbert, M. D. and Barandiaran, X. E. (2014). Modeling habits as self-sustaining patterns of sensorimotor behavior. *Frontiers in Human Neuroscience*, 8(August):1–15.

Maye, A. and Engel, A. (2011). A discrete computational model of sensorimotor contingencies for object perception and control of behavior. In *IEEE International Conference on Robotics and Automation*, pages 3810–3815.

O'Regan, J. K. and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5):939–973.

Woolford, F. M. G. and Egbert, M. D. (2019). Behavioural variety of a node-based sensorimotor-to-motor map. *Adaptive Behavior*.

Zarco, M. and Egbert, M. (2019). Different forms of random motor activity scaffold the formation of different habits in a simulated robot. *Artificial Life Conference Proceedings*, (31):582–589.