Habit-based Regulation of Essential Variables

Matthew Egbert[1] and Lola Cañamero[1]

[1] Embodied Emotion, Cognition and (Inter-)Action Lab,
University of Hertfordshire, School of Computer Science, Hatfield, UK, AL10 9AB
mde@matthewegbert.com

Abstract

A variety of models have been developed to investigate "homeostatic adaptation," a mechanism inspired by Ashby's homeostat, where a plastic control medium is reorganized until one or more essential variables are maintained within predefined limits. In these models, "habits" emerge, defined as behavior-generating mechanisms that rely upon their own influence to maintain the conditions necessary for their own persistence. In this paper, we present a recently developed sensorimotor-habit-based controller that is coupled to a simulated two-wheeled robot with a simulated metabolism. The simulation is used to demonstrate how habits can have the same essential variable(s) as the metabolic or "biological" organism that is performing the behavior, and that in certain conditions when this is the case, the emergent habits will tend to stabilize essential variables within viability limits. The model also demonstrates that an explicit pre-specification of (A) which variables should induce plasticity and (B) which values of those variables should induce plasticity is not always necessary for homeostatic adaptation of behavior.

Introduction

Over the past few decades, a range of "embodied" approaches to the study of the mind have emerged, including: enaction (Stewart et al., 2010), dynamical approaches to cognition (Beer, 1997), the sensorimotor approach to perception (O'Regan and Noë, 2001), and embodied AI/Robotics (Brooks, 1990; Pfeifer and Bongard, 2006). Unifying these approaches is the notion that instead of investigating abstract computational problems such as chess, research in robotics and AI should instead focus upon embodied, embedded agents that use ongoing interactions with their environment to accomplish "intelligent behavior," where intelligence is evaluated by its contribution to the survival of the agent, (see e. g. Steels, 1995). But, the notion of system survival or "viability" in the context of robotics can be more challenging than it first appears. Biological organisms are far-from-equilibrium, dissipative structures that depend upon ongoing processes of self-production and self-maintenance to counteract their tendency to degrade (Kauffman, 2000; Nicolis and Prigogine, 1977; Schrödinger, 1944; Maturana and Varela, 1980), but conventional robots do not share this property and thus there can be challenges in relating the behavior of artifacts to the behavior of biological organisms. This difficulty has led to different emphasis placed upon the metabolic organization within the embodiment community, with some proponents more or less ignoring the concept and others arguing that the metabolic organization is fundamental (see Froese and Ziemke, 2009). As an example of the former, metabolism is seen as largely irrelevant to the sensorimotor approach to perception, with suggestions that a missile guidance system could have some minimal perceptions (O'Regan and Noë, 2001, p. 82) and that a machine, if sufficiently complex, could be considered aware (with no mention of self-production). In contrast, many studies in embodied- and evolutionary-robotics acknowledge the metabolic-organization and its role of defining viability limits as important and set out to investigate how regulation can be accomplished that maintains viability. Most embodied robotics studies do not model in detail the processes through which the viability-limits are defined (e. g. metabolism), as the focus is instead upon how a particular form of regulation can be accomplished (e. g. Avila-Garcia and Cañamero, 2004; Beer, 1995), but some of our recent models include both behavioral and the metabolic dynamics that define the viability limits, demonstrating how organisms can respond to the processes that define their viability limits to perform an ongoing, in-the-moment evaluation of the environment (Egbert et al., 2009), that allows an organism to behave in a history-dependent and adaptive way that integrates multiple environmental and internal factors into an adaptive behavior (Egbert et al., 2010b; Egbert, 2013) that may play an interesting role of facilitating adaptive evolution (Egbert et al., 2010a).

Some "enactivists" (Stewart et al., 2010) believe that the metabolic-organization fundamentally underlies mind. In particular, it has been argued that the metabolic-organization underlies a modern concept of intrinsic teleology, in which an organism's survival grounds teleological (Weber and Varela, 2002) and normative (Barandiaran and Egbert, 2013) descriptions of its behavior. In this view, organisms act to satisfy the needs that are the result of their own precarious,

metabolic organization (Weber and Varela, 2002). But of course, not all behavior, adaptive or otherwise, is a direct response to biological needs and dynamics; not all behavior occurs to maintain biological essential variables within viability limits, and so what can biological embodiment contribute, to "higher" forms of adaptive behavior, i. e. beyond the metabolism-based chemotaxis of bacteria? One proposal stems from the observation that aspects of perception, learning and sensorimotor behavior appear to have properties in common with self-maintaining organisms. In particular, an analogy is drawn between *habits* seen as self-maintaining patterns of behavior, and *biological organisms* seen as self-maintaining metabolic entities (Di Paolo, 2003; Barandiaran, 2007, 2008). By viewing habits as having the same precarious, self-maintaining organization as biological entities, it becomes possible to understand how, in a way similar to the grounding of many survival-based behaviors in metabolism, other behaviors can be grounded in the needs of the habits themselves. In metabolism-based chemotaxis, bacteria act to maintain their biological essential variables and analogously, habit-based behavior maintains the conditions necessary for the persistence of the behavior itself. Or, put another way, biological embodiment grounds survival, the "mother of all values" (Weber and Varela, 2002; Di Paolo, 2003), and habits are autonomous "mental life forms," that ground the others (Di Paolo, 2003; Barandiaran, 2007, 2008).

In a set of minimalistic computational models, Di Paolo (2000, 2003) demonstrated how a behavior-generating mechanism can stabilize the conditions necessary for its own persistence. The first model involved a simple simulated two-wheel, two-light-sensor robot, controlled by a plastic continuous time recurrent neural network. In a design inspired by Ashby's homeostat (Ashby, 1952), the weights of the connections in the neural network were fixed when the neural firing-rate was within a predefined region, but outside of this region the weights changed according to deterministic plasticity rules. A genetic algorithm succeeded at identifying model parameters (initial weights, which plasticity rules applied to which synapses, etc.), that resulted in networks that both avoid plasticity and perform phototaxis. We can say that the neural network configurations that emerged in this model were self-maintaining behavior-generating mechanisms. Why? A configuration of weights in the neural network is precarious in the sense that it can only persist if plasticity is avoided; it has viability-limits in that neuron firing rates must not leave predefined bounds if the behavior-generating mechanism is to persist. The firing rates of the neurons are influenced by internal neuron activity, but also by the sensorimotor dynamics of the behavior: the neurons influence the motors which influence how the environment influences the sensors, which, coming full circle, influence the neurons. The stability of a particular configuration weights of the NN therefore depends, in part, upon the behavior that it drives, and thus can be considered to be a self-maintaining behavior-generating mechanism.

In this first model of Di Paolo's, there is no modeled metabolism, biological essential variables or biological viability limits. The behavior is phototaxis because of selection pressures applied during the artificial evolution. We could imagine that the agent has a biological (i. e. metabolic) need to move toward the light, but even so, the coupling of this need and the self-maintaining behavioral mechanism was accomplished through an evolutionary process. Is there a way that the self-maintaining behavior-generating mechanisms (henceforth "habits") and the self-maintenance of the organism could be more tightly integrated? In order to address this, Di Paolo (2003) developed a second model consisting of a simulated, Braitenberg-inspired robot, with plastic mappings between sensors and motors. These mappings undergo stochastic change when the battery (an analog to metabolism) is outside of certain pre-defined limits. This is a step forward in that the biological needs and the habits are intertwined in a cycle of dependence, where the maintenance of a viable battery level depends upon the habit, but also the stability of the habit depends upon the battery level remaining within bounds, for if the habit drives the metabolism outside of the biological viability limits, the behavior-generating mechanism will also lose viability due to the stochastic weight-change process. In a different approach to coupling behavior to metabolic needs, Iizuka et al. (2013) selected the limits for the induction of plasticity such that when there was no sensor activity, the weights of the network would change. In this way, the system could only become stable when it was performing phototaxis. The relationship between the essential variable (a "photosynthetic metabolism" that justifies phototaxis as a target behavior) and the behavior has been hard-wired by the designer, in the sense that the robot can adapt to inversions of its visual field, but could not adapt to its phototactic-needs becoming photophobic-needs.

For all of the models presented above, the limits of plasticity, i. e. the surface between the states where plasticity does occur and where it does not, are parameters of the model that are pre-specified. It is also the case that for every model, one or more variables are given a special privileged status as a "plasticity inducing variable". In the first model of Di Paolo's presented above, and in Iizuka's model, these are the neuron firing rates. In the second model of Di Paolo's, it is the state of the "biological" essential variable, i. e. the battery state, and it is through this explicit specialization of this variable that the biological dynamics and the mental dynamics are coupled. In this paper, we use a newly developed model of habits, coupled to a simulated two-wheeled robot with simulated metabolic dynamics to demonstrate how it can be possible to couple self-maintaining behavior-generating mechanisms with biological essential variables without pre-specifying the region of

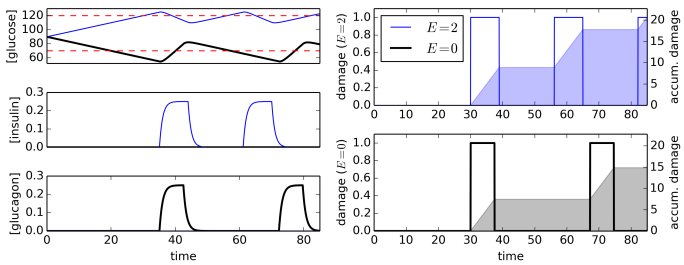

Figure 1: Essential variable dynamics in two worst case scenarios, always feeding (blue) and never feeding (black). Left: Glucose, insulin and glucagon dynamics. Viability limits for the essential variables are indicated by dashed red lines. Right: damage and accumulated damage.

plasticity and moreover, without pre-specifying which variables are plasticity-inducing variables.

In the next section we describe the robot, its simulated metabolism and its habit-based controller. We then describe two scenarios that we use to demonstrate that by including the state of the metabolism as a sensory-variable, self-reinforcing patterns of behavior emerge that stabilize that metabolic dynamics. We present the results of simulations of these scenarios before interpreting and discussing them.

Simulation

Robot and environment. We simulate an simple robotic agent embedded in a two-dimensional square environment 8 units wide, with periodic boundary conditions. The robot has two directional light sensors and two independently driven motorized wheels. The motion of the robot is $\dot{x} = \cos(\alpha)(m_l + m_r)$; $\dot{y} = \sin(\alpha)(m_l + m_r)$; $\dot{\alpha} = 2(m_r - m_l)$, where x,y is the robots spatial position, $\alpha \in [-\pi, \pi]$ is the robots orientation and $m_l \in [-1, 1]$ and $m_r \in [-1, 1]$ are the robots left and right motor speeds. The robot's light sensors are located at $x + r \cdot \cos(\alpha + \beta)$,$y + r \cdot \sin(\alpha + \beta)$, where $r = 0.25$ is the robot's radius and $\beta = \pm\pi/3$ is the angular offset of the sensors from α, the heading of the robot. The activation of each sensor is determined by the equation below, where $\boldsymbol{b} = [\cos(\alpha + \beta), \sin(\alpha + \beta)]$ is a unit vector indicating the direction that the sensor is facing, $\boldsymbol{c}$ is the vector from the sensor to the light (which is located in the center of the arena at $(0, 0)$), and D is the distance from the sensor to the light.

$$s = \frac{(\boldsymbol{b} \cdot ||\boldsymbol{c}||)^+}{1 + D^2} \tag{1}$$

Metabolism. We simulate the robot as having a metabolism with intrinsic dynamics that are indirectly influenced by the motor behavior. The simulated metabolism

param.	**value**	**description**
c	1.0	rate of glucose use / consumption
f_U, f_I	$20, 0.1$	efficiency of hormonal modulation
b_U, b_I	$60.0, 120.0$	threshold G-concentrations
c_U, c_I	$0.25, 0.25$	rate of hormonal production
d_U, d_I	$1.0, 1.0$	rate of hormone elimination
τ	5.0	delay in hormonal response

Table 1: Metabolism-related parameters

is inspired by the blood-sugar dynamics of a diabetic, where hormonal regulation prevents the blood-sugar from diverging, but is insufficient to prevent blood-glucose levels from leaving healthy limits. The dynamics of diabetes has been simulated since this work is part of the ALIZ-E project, where we are investigating how to help diabetic children learn to manage their disease. Part of this support involves developing a better understanding of how behaviors relate to essential variables such as blood-sugar, as well as to reflexive hormonal modulators such as insulin and glucagon (Lewis and Cañamero, 2014), and how diabetes-related (good and bad) habits form and how they can be changed (from bad to good) in support of self-efficacy.

The modeled metabolism is not intended to be a realistic simulation of blood-sugar dynamics, but just to qualitatively capture the dynamics of an hormone regulated essential variable that is inadequately regulated, i. e. a variable for which the hormonal regulation is insufficient to keep the system within its viability limits. The model consists of three coupled delayed differential equations, which represent: G, blood-glucose concentration, the essential variable which must remain within limits if the system is to be considered in a healthy state; I, the concentration of insulin, a hormone that removes G from the blood when it is above a threshold, and U, the concentration of glucagon, a hormone that releases G into the blood when it is below a threshold. In these equations, the function $[a < b] \equiv 1$ when $a < b$, and 0 otherwise.

$$\frac{\mathrm{d}G}{\mathrm{dt}} = E + f_U U - f_I I G - c \tag{2}$$

$$\frac{\mathrm{d}I}{\mathrm{dt}} = [b_I < G_{t-\tau}] c_I - d_I I \tag{3}$$

$$\frac{\mathrm{d}U}{\mathrm{dt}} = [G_{t-\tau} < b_U] c_U - d_U U \tag{4}$$

We defined viability limits such that the system is considered to be healthy if $G \in [b_U, b_I]$. Leaving the viability limits will eventually trigger hormonal regulation of G back into the viability region (with a delay of τ), but the model is configured such that similar to a diabetic, the hormonal regulation is insufficient to maintain G within healthy limits. When the robot is within 2 spatial units of the light it is considered to be "feeding" and the variable E is set to 2, and otherwise $E = 0$. Thus, the behavior of the robot

influences G and, as we shall see, if certain patterns of behavior are performed, it is possible for the value of G to remain within the viability limits indefinitely. However, if the robot performs non-ideal behaviors, G will leave the viability limits, and thereby accumulate "damage," defined as the quantity of time when $G \notin [b_U, b_I]$. Figure 1 shows example trajectories for G, I and U for two worst-case behaviors (the robot always eating, and the robot never eating) with the plots on the right indicating the damage and accumulated damage suffered in these scenarios. Throughout this paper, time and time-related values such as τ are specified in arbitrary time-units, where one time-unit is the amount of time that a robot traveling at full-speed moves 2 spatial units.

Symbol	Description
$\boldsymbol{x}$	current SM-state
$N_{\boldsymbol{p}}$	SM-state associated with node N (in normalized SM-space coordinates)
$N_{\boldsymbol{v}}$	SM-velocity indicated by node N (in normalized SM-space coordinates)
N_w	weight of node N
$d(\boldsymbol{x}, \boldsymbol{y})$	distance function between two SM-states
$\omega(N_w)$	function describing how the weight of a node scales its influence
$\phi(\boldsymbol{y})$	function describing the "familiarity" (local density of nodes) of SM-state $\boldsymbol{y}$

Table 2: Glossary of symbols and brief descriptions.

Habit-based controller We have recently developed a plastic, self-modifying dynamical system called an Iterant Deformable Sensorimotor Medium (IDSM) (Egbert and Barandiaran, 2014). This system was designed to act as a robot controller that supports the formation of "habits" conceived of as precarious, self-maintaining patterns of sensorimotor behavior. When coupled to a robot's sensors and motors, the IDSM (1) causes the robot to repeat behaviors that it has performed in the past, and (2) allows for the reinforcement of patterns of behavior through repetition, such that the more frequently and recently a pattern of behavior has been performed, the more likely it is to be performed again in the future. If a pattern of behavior is not performed for a period of time, it becomes less likely to be re-enacted, but when behaviors are performed, they become more likely to be repeated in the future, and in this way, self-maintaining patterns of behavior emerge. Metaphorically, the IDSM works similarly to the paths made by animals through the woods or through a field of grass. As sensorimotor trajectories are experienced, pathways are worn in to the IDSM's "sensorimotor-space," such that future sensorimotor pathways are likely to be similar to those pathways that have been taken in the past. In the remainder of this section, we provide an overview of our IDSM architecture. Much of the text here comes from (Egbert and Barandiaran, 2014), which provides a much more detailed description of the IDSM and its dynamics.

The IDSM operates by developing and maintaining a history of sensorimotor (SM) dynamics. This history takes form of many *nodes*, where each node describes the SM-velocity at a SM-state at some point in the past. As the agent behaves, and its SM-state changes, nodes are added, such that a record is constructed of how sensors and motors have changed for various SM-states during the system's history. These are used to determine future motor-actions such that when a familiar SM-state is encountered, the IDSM produces behavior that is similar to the behavior that was performed when the agent was in a similar situation in the past.

More formally, each node is a tuple of two vectors and a scalar, $N = \langle \boldsymbol{p}, \boldsymbol{v}, w \rangle$, where $\boldsymbol{p}$ indicates the SM-state associated with the node (referred to as the node's "position" in SM-space), $\boldsymbol{v}$ indicates a velocity of SM-change, and the scalar, w indicates the "weight" of the node, a value that partially determines the overall influence of the node. We shall refer to these components using a subscript notation, where the position, SM-velocity, and weight of node N are written as $N_{\boldsymbol{p}}$ and $N_{\boldsymbol{v}}$ and N_w, respectively.

As a robot controlled by the IDSM moves through SM-states, new nodes are created recording the SM-velocities experienced at different SM-states. Specifically, when a new node is created, its "position," $N_{\boldsymbol{p}}$ is set to the current SM-state; its "velocity," $N_{\boldsymbol{v}}$ is set to the current rate of change in each SM-dimension, and its weight, N_w is set to 0 (an initial value that does not imply that the node is ineffectual, see below). The two vector terms ($N_{\boldsymbol{p}}$ and $N_{\boldsymbol{v}}$) are calculated in a normalized sensorimotor space, where the range of all sensor and motor values are linearly scaled to lie, in each dimension, between 0 and 1.

New nodes are only added when the density of nodes near the current SM-state, as described by the function ϕ, is less than a threshold value, $\phi(\boldsymbol{x}) < k_t = 1$. This density function, ϕ, can be thought of as a measure of how many nodes there are near to the SM-state $\boldsymbol{x}$, and how heavily weighted those nodes are. Loosely speaking, it is a measure of how "familiar" the SM-state is, and it is calculated by summing a non-linear function of the distance from every node to the current SM-state, $d(N_{\boldsymbol{p}}, \boldsymbol{x})$, scaled by a sigmoidal function of the node's weight, $\omega(N_w)$, as described in Equations 5–7.

$$\phi(x) = \sum_N \omega(N_w) \cdot d(N_{\boldsymbol{p}}, \boldsymbol{x}) \tag{5}$$

$$\omega(N_w) = \frac{2}{1 + \exp(-k_\omega N_w)};\ k_\omega = 0.025 \tag{6}$$

$$d(N_{\boldsymbol{p}}, x) = \frac{2}{1 + \exp(k_d ||N_{\boldsymbol{p}} - \boldsymbol{x}||^2)};\ k_d = 1000 \tag{7}$$

After a node is created, its weight changes according to equation 8, where the first term represents a steady degrada-

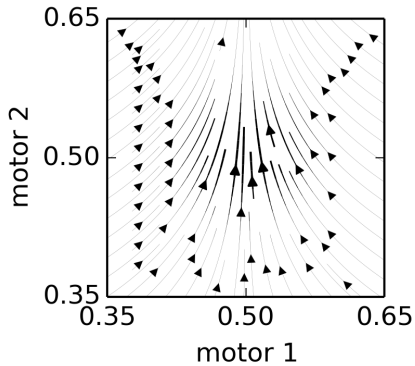

Figure 2: The influence of a single node, with $N_{\boldsymbol{p}} = (0.5, 0.5)$, $N_{\boldsymbol{v}} = (0, 0.1)$, and $N_w = 0$. In this didactic scenario, $N_{\boldsymbol{v}}$ only has a non-zero component in the motor-2 dimension, and thus the node's "velocity" influence causes motor-2 to increase, and its "attraction" influence causes motor-1 to approach a state of 0.5.

tion of the node's influence, and the second term represents a strengthening of the node that occurs when the current SM-state is close to the node's position. This latter term allows for the self-reinforcement/self-maintenance of patterns of behavior, such that patterns of behavior that are repeated are more likely to persist than those that are not reinforced.

$$\frac{\mathrm{d}N_w}{\mathrm{d}t} = -1.0 + r(N, \boldsymbol{x}) \tag{8}$$

$$r(N, \boldsymbol{x}) = kd(N_{\boldsymbol{p}}, \boldsymbol{x}) \tag{9}$$

A short period of time after creation (10 simulated time-units), nodes are activated, meaning that they are added to the pool of nodes that influence the motor state. Every activated node influences the motor state, but at any one time only a subset of these will have a substantial influence, for the influence of a node is scaled non-linearly by its distance from the current SM-state by the same distance function used in ϕ above (Equation 7). The influence of each node is also scaled by its weight according to Equation 6, and thus nodes that are close to the current SM-state and nodes with higher weights have a greater influence.

The influence of a node can be broken down into two factors: a "velocity" factor and an "attraction" factor. The velocity factor is simply the motor components of the $N_{\boldsymbol{v}}$ vector, but the attraction factor, is slightly more complicated. It is a sensorimotor-"force" that draws the system towards the node. This tends to result in a motion in SM-space towards regions of SM-space that are familiar, i.e. for which there is a higher density of nodes and it can compensate for stochasticity in the environment or perturbations to behavior (see Egbert and Barandiaran, 2014 for details). The attraction vector has its component parallel to $N_{\boldsymbol{v}}$ removed to prevent it from interfering with the velocity influence of the node (again, see Egbert and Barandiaran, 2014 for details). Figure 2 provides a visualization of the influence of a single node in a hypothetical 2-motor, 0-sensor IDSM. In this example, $N_{\boldsymbol{v}}$ is exactly vertical, so all horizontal motion is due to the attraction component, and vertical motion is due to the velocity component.

Equations 10 – 11 describe how the IDSM influences the motor state. The velocity and attraction influences of every node are scaled by the node's weight and distance to the SM-state, and then these are all summed before being scaled by the density of the nodes at the current SM-state such that the influence of all the nodes is averaged and not cumulative. Obviously, the IDSM only has direct control of its motors and the sensor-components of the SM-state are determined by the systems interaction with its environment. Accordingly, the superscripted-$\boldsymbol{\mu}$ notation in the equations below indicates where we are only using the motor-components of the indicated vector terms.

$$\frac{\mathrm{d}\boldsymbol{\mu}}{\mathrm{d}t} = \frac{1}{\phi(\boldsymbol{x})} \sum_N \omega(N_w) \cdot d(N_{\boldsymbol{p}}, \boldsymbol{x}) \cdot (\underbrace{N_{\boldsymbol{v}}}_{\text{Vel.}} + \underbrace{A(N_{\boldsymbol{p}} - \boldsymbol{x}, N_{\boldsymbol{v}})}_{\text{Attraction}})^{\boldsymbol{\mu}} \tag{10}$$

$$A(\boldsymbol{a}, N_{\boldsymbol{v}}) = \boldsymbol{a} - \boldsymbol{a} \cdot \frac{N_{\boldsymbol{v}}}{||N_{\boldsymbol{v}}||} \tag{11}$$

Experiment and Control Scenarios We compare two scenarios. In the experimental scenario, the SM-space of the IDSM has two motor dimensions: (the left and right motor of the robot) and three sensory dimensions: its two directional light-sensors, and a direct sensory perception of its essential variable, G (scaled linearly such that the range $G \in [50, 130]$ lies in normalized sensorimotor coordinates in $[0, 1]$). In the control scenario, everything is the same except that the IDSM is not sensitive to G. To keep the total number of SM-dimensions the same in both scenarios, the sensitivity to G was replaced with a motor that has no effect whatsoever. The control is included primarily to show that the task is not trivially solved.

We simulated 25 trials of each scenario. At the start of each trial, we randomly initialized the IDSM with 10000 nodes. These were generated by performing 200 random walks in the 5 dimensional SM-space, each starting from a random location within the SM-space with subsequent loci calculated according to the following equation, $\boldsymbol{l}_{i+1} = \boldsymbol{l}_i + \boldsymbol{r}$, where the components of $\boldsymbol{r}$ are selected from a flat distribution $[-0.05, 0.05]$ and where any components that would take l_i out of the normalized SM-volume are inverted. Nodes were added at each locus of the walk l_i with $N_{\boldsymbol{p}}$ set to l_i, $N_{\boldsymbol{v}}$ set to $l_{i+1} - l_i$, and $N_w = 0$. We then placed the robot at a random initial location within the arena, with the essential variable initialized to a value at the center of its viability region ($G = 90$) and the concentration of the regulatory hormones I and U set to 0. We then allowed

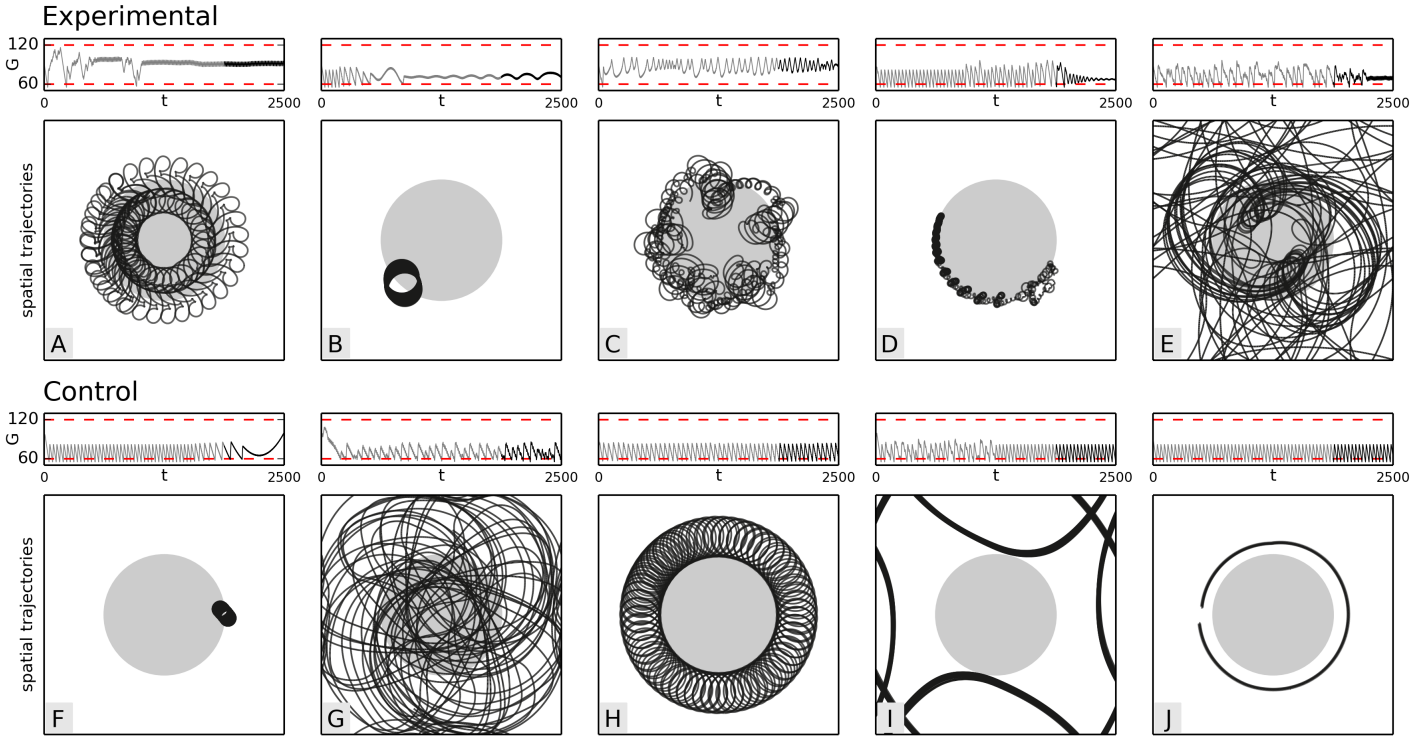

Figure 3: Essential variable and spatial trajectories of the five best performing control and experimental agents. The square plots indicate the spatial trajectories taken by each agent during the final 625 time-units of its simulation (the period during which damage was evaluated). In these plots, the filled-circle indicates the feeding-region. Above each spatial plot, the essential variable, G is plotted against time, with the viability-limits indicated by red-dashed lines and the period corresponding to the spatial plot indicated by a darker line.

the IDSM to control the robot and tracked the position of the robot and the trajectories of G, I, and U for 2500 simulated time units, so that we could evaluate the behavior of the robot, and the extent of its success at maintaining the essential variable within viability limits.

Results

To evaluate the performance of each trial, we measured the amount of damage accumulated during the final 625 time-units of each simulation. Figure 3 shows the spatial trajectories and the glucose concentration trajectory plotted against time for the top 5 performing experimental and control trials. We can see that by the end of the simulation, all five of the plotted experimental trials have behavior that maintains G within these limits (although the trajectory in Trial B appears to be on an amplitude-increasing cycle that may eventually leave the viability limits). In comparison, none of the control trials appear to have stabilized G within the viability limits. Three of the experimental agents manage to avoid incurring any damage during this period, and none of the control agents are as successful. A variety of spatial trajectories can be observed, both in the spatial dynamics and in the dynamics of G. Figure 4 shows the accumulated damage for each trial during this last quarter of the simulation. A Mann-Whitney-Wilcoxon test of these values indicates that the experimental agents are better at maintaining the essential variable within limits to a statistically significant degree ($z = -3.13, p < 0.002$).

Discussion

The IDSM supports the formation of self-maintaining patterns of behavior by (1) assembling a collection of "nodes" that track the SM-state-velocity for different SM-states, (2) using these nodes to drive later behavior, and (3) having these nodes, which perpetually degrade, depend upon a mechanism of self-reinforcement to persist. The self-reinforcement of a node is accomplished by the re-visitation of SM-states near to the node's "position" ($N_{\boldsymbol{p}}$), and so only patterns of behavior that repeatedly visit SM-states can persist. Therefore, in the experimental scenario where the IDSM is sensitive to the essential variable, G, the only patterns of behavior that will persist will be those where values of G are regularly revisited in a way that correlates with the other sensorimotor state variables. In several of the experi-

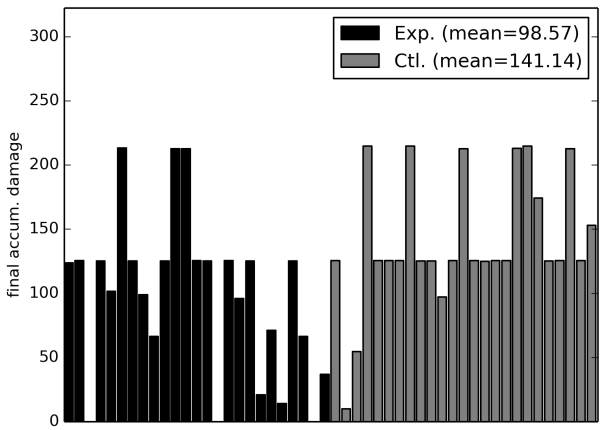

Figure 4: Final accumulated damage for each of 25 experimental and 25 control trials. A higher bar indicates greater damage incurred, i. e. worse performance.

mental trials, the stable pattern that emerged was one where G was within the viability limits. Why? One possible explanation is that the dynamics of G are more reliable within the viable region than near or outside of it. When G has been inside the viable region for the system's recent history, $U = 0$ and $I = 0$, and thus G changes at a constant rate determined by E. In contrast, when G has recently left the viability limits, one of the hormones will increase in concentration and the way that G changes becomes less correlated with the other sensorimotor variables, (especially given the delay in the differential equations). Thus it seems more likely for the system to find a repeating sensorimotor pattern when it is within the viable region.

For Ashby's homeostat and the robot controllers that it inspired, the stability of the controller has been related to the notion of an *ecological invariant*; a relationship with the environment that is maintained by the behavior and that the behavior depends upon. By including the state of G as an interoceptive sensory input, the dynamics of G become part of the habit's sensorimotor "environment" (Buhrmann et al., 2013), such that a behavior must cause G to change in a repetitive way if it is to persist. G is the biological essential variable, but once it is added as a sensory-variable to the IDSM, its dynamics are also essential to the persistence of the habit. The mental and biological autonomous structures are thus intertwined in the sense that they share an essential variable.

Perhaps ecological *invariant* is a bit of a misnomer in this model, as it is not the fixed state of G or any other variable that determines the stability of a pattern of behavior, but more the reliability of a repeated dynamic. In both the control and experimental agents, the behaviors displayed towards the end of the trials are cyclical. This is due largely to the dynamics of the IDSM, where again, only patterns of SM-activity that repeat are reinforced, and as discussed above, only reliable interaction with the environment can result in repeated patterns of SM-activity. At times, the internal dynamics and the environmental interaction are "discordant", in the sense that the motor activity driven by the IDSM does not result in reliable sensory input and the internal dynamics do not "resonate" with the environment in a self-stabilizing manner. As an example of this, consider the more chaotic behavior in Trial E of Figure 3, where the agent is moving around the whole arena, irregularly encountering the feeding area. These irregular sensorimotor trajectories are inherently less stable than those that cause a repeated pattern of sensorimotor state, such as those demonstrated by the subsequent, radially-symmetrical patterns in Trial E, and most of the other agents depicted in Figure 3 (perhaps most apparent in trials A, H, I and J).

In previous attempts to couple biological essential variables to self-maintaining behavior-generating mechanisms, it has been argued that it is necessary to have two nested closed-loops; the first loop being a behavioral coupling between the environment and the organism and the second being an evaluation of the first via an essential variable, such that when the biological essential variable goes out of bounds, the behavior-generating mechanism is reorganized (Di Paolo, 2003). The IDSM-based habits in the experimental scenario are indeed dependent upon the maintenance of a biological essential variable, but we would argue that the two-nested-feedback loop description is not the best way to describe the homeostatic adaptation demonstrated in our model, in that the relationship between the operating limits of the habits and the operating limits of the simulated metabolism are more integrated here than in previous models. The habit does not depend upon the behavior because of a prescribed threshold and response; i. e. it is not due to a random reorganization of the system that is brought about by a pre-specified essential variable going outside of some pre-specified viability limits. Instead, the stability of the behavior and its behavior-generating mechanism is directly dependent upon the repetition of a particular trajectory of the sensorimotor variables, including G. We propose that in this paper we see an example of homeostatic adaptation that blends these two feedback loops into one, suggesting that having two nested feedback loops may not always be necessary.

The system does not always find stability within the viable region, and indeed in several cases (about half) the experimental agents fare no better than the control agents. In these cases, habits have still emerged, but they are unhealthy habits in the sense that they do not maintain the biological essential variable within limits. In ongoing work, and as part of the ALIZ-E project, we are working with our colleagues to develop "diabetic robots" with a simulated glucose/insulin/glucagon metabolism, that diabetic children can interact with in different ways to investigate how inter-

action, or modification of the environment could modulate unhealthy habits into healthy habits. By helping the robot transform unhealthy habits into healthy habits, the diabetic children will develop greater self-efficacy, self-confidence and self-esteem, enabling them to better manage their disease (Lewis and Cañamero, 2014). We believe that models such as that presented here can provide insight and fresh perspectives into the relationship between habits and health and how such habits can be better managed.

Conclusion

We have presented our most recent exploration with the IDSM, demonstrating how it can regulate behavior to stabilize essential variables within limits simply by having the state of the essential variable included among its sensors. When habits emerge in this configuration, the habit itself depends upon the dynamics of the biological essential variable, and in this way have demonstrated the possibility of more tightly integrated biological and mental autonomous structures.

Acknowledgments

This research was funded by the European Commission as part of the ALIZ-E project (FP7-ICT-248116). The opinions expressed are solely the authors. The authors would like to thank Tom Froese, Matt Lewis and members of the Embodied Emotion, Cognition and (Inter-)Action Lab for discussions of the research presented above.

References

Ashby, W. R. (1952). *Design for a Brain: the origin of adaptative behaviour*. J. Wiley, London, 2nd edition.

Avila-Garcia, O. and Cañamero, L. (2004). Using hormonal feedback to modulate action selection in a competitive scenario. In *Proc. 8th Intl. Conf. on Simulation of Adaptive Behavior (SAB 2004)*, pages 243–252, Cambridge, MA. MIT Press.

Barandiaran, X. (2007). Mental life: conceptual models and sythetic methodologies for a post-cognitivist psychology. In Wallace, B., Ross, A., Anderson, T., and Davies, J., editors, *The World, the Mind and the Body: Psychology after cognitivism*, pages 49—90. Imprint Academic.

Barandiaran, X. E. (2008). *Mental Life. A naturalized approach to the autonomy of cognitive agents*. PhD thesis, University of the Basque Country, Spain.

Barandiaran, X. E. and Egbert, M. D. (2013). Norm-establishing and norm-following in autonomous agency. *Artif. Life.*, pages 1–24.

Beer, R. D. (1995). On the dynamics of small continuous-time recurrent neural networks. *Adapt. Behav.*, 3(4):469–509.

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

Brooks, R. A. (1990). Elephants don't play chess. *Robotics and Autonomous Systems*, 6(1-2):3–15.

Buhrmann, T., Paolo, E. A. D., and Barandiaran, X. (2013). A dynamical systems account of sensorimotor contingencies. *Front. Psychol.*, 4:285.

Di Paolo, E. A. (2000). Homeostatic adaptation to inversion of the visual field and other sensorimotor disruptions. In *From animals to animats 6: Proceedings of the 6th International Conference on the Simulation of Adaptive Behavior*, page 440–449.

Di Paolo, E. A. (2003). Organismically-inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop. In Murase, K. and Asakura, T., editors, *Dynamical systems approach to embodiment and sociality*, pages 19—42. Advanced Knowledge International, Adelaide, Australia.

Egbert, M. D. (2013). Bacterial chemotaxis: Introverted or extroverted? A comparison of the advantages and disadvantages of basic forms of metabolism-based and metabolism-independent behavior using a computational model. *PLoS ONE*, 8(5):e63617.

Egbert, M. D. and Barandiaran, X. E. (2014). Modelling habits as self-sustaining patterns of sensorimotor behavior. *Submitted to Frontiers in Human Neuroscience*.

Egbert, M. D., Barandiaran, X. E., and Di Paolo, E. A. (2010a). Behavioral metabolution: Metabolism based behavior enables new forms of adaptation and evolution. In *Artificial Life XII Proceedings of the Twelfth International Conference on the Simulation and Synthesis of Living Systems*, pages 213—220. MIT Press.

Egbert, M. D., Barandiaran, X. E., and Di Paolo, E. A. (2010b). A minimal model of metabolism-based chemotaxis. *PLoS Comput Biol*, 6(12):e1001004.

Egbert, M. D., Di Paolo, E. A., and Barandiaran, X. E. (2009). Chemo-ethology of an adaptive protocell: Sensorless sensitivity to implicit viability conditions. In *Advances in Artificial Life, Proceedings of the 10th European Conference on Artificial Life, ECAL.*, pages 242—250. Springer.

Froese, T. and Ziemke, T. (2009). Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence*, 173(3–4):466–500.

Iizuka, H., Ando, H., and Maeda, T. (2013). Extended homeostatic adaptation model with metabolic causation in plasticity mechanism—toward constructing a dynamic neural network model for mental imagery. *Adapt. Behav.*, page 1059712313488426.

Kauffman, S. (2000). *Investigations*. Oxford University Press, USA.

Lewis, M. and Cañamero, L. (2014). An affective autonomous robot toddler to support the development of self-efficacy in diabetic children. In *Proc. 23rd IEEE Intl. Symp. on Human-Robot Interactive Communication (IEEE RO-MAN 2014)*.

Maturana, H. R. and Varela, F. J. (1980). *Autopoiesis and cognition: the realization of the living*. Springer.

Nicolis, G. and Prigogine, I. (1977). *Self-organization in nonequilibrium systems: from dissipative structures to order through fluctuations*. Wiley, New York, NY.

O'Regan, J. K. and Noë, A. (2001). What it is like to see: A sensorimotor theory of perceptual experience. *Synthese*, 129(1):79–103.

Pfeifer, R. and Bongard, J. C. (2006). *How the Body Shapes the Way We Think: A New View of Intelligence*. The MIT Press.

Schrödinger, E. (1944). *What is life?: The physical aspect of the living cell*. The University Press.

Steels, L. (1995). Intelligence — dynamics and representations. In Steels, L., editor, *The Biology and Technology of Intelligent Autonomous Agents*, number 144 in NATO ASI Series, pages 72–89. Springer, Berlin.

Stewart, J., Gapenne, O., and Di Paolo, E. A. (2010). *Enaction: Toward a New Paradigm for Cognitive Science*. The MIT Press.

Weber, A. and Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1(2):97–125.