

# A Hormone-Driven Epigenetic Mechanism for Adaptation in Autonomous Robots

John Lones, Matthew Lewis, Lola Cañamero  
Embodied Emotion, Cognition and (Inter-)Action Lab  
School of Computer Science, University of Hertfordshire  
Hatfield, Herts AL10 9AB, U.K.  
{J.Lones;M.Lewis4;L.Canamero}@herts.ac.uk

**Abstract**—Different epigenetic mechanisms provide biological organisms with the ability to adjust their physiology and/or morphology and adapt to a wide range of challenges posed by their environments. In particular, one type of epigenetic processes in which hormone concentrations are linked to the regulation of hormone receptors has been shown to have implications for behavioral development. In this paper, taking inspiration from these biological processes, we investigate whether an epigenetic model based on the concept of hormonal regulation of receptors can provide a similarly robust and general adaptive mechanism for autonomous robots. We have implemented our model using a Koala robot, and tested it in a series of experiments in six different environments with varying challenges to negotiate. Our results, including the emergence of varied behaviors that permit the robot to exploit its current environment, demonstrate the potential of our epigenetic model as a general mechanism for adaptation in autonomous robots.

**Index Terms**—Autonomous Robot, Epigenetic development, Hormonal modulation, Emergent behavior, Allostasis.

## 1 INTRODUCTION

ONE of the biggest challenges in autonomous robot research is developing controllers that can adapt to the unstructured, novel and dynamic conditions found in the real world. As pointed out by Krichmar [1], while significant advances have been seen in the capabilities of autonomous systems, controllers tend to be based on mechanisms that are developed to operate in highly task-specific and relatively controlled environments. While these models have proven to foster adaptation in their specific environments, questions must be asked about their ability to recreate this performance in less controlled settings, where even small changes in conditions can have significant, and often unpredictable, implications for the behavior of robots [2], [3], [4]. To operate and adapt autonomously in real world environments, robot controllers must not only be robust, but also have flexible and general adaptation capabilities, allowing a robot to adapt to different environmental conditions changing across environments and over time. In biological organisms, epigenetic processes [5], [6], [7], [8] provide such an adaptation mechanism. Drawing inspiration from development in biological organisms, we claim that an epigenetic process in which environmental stimuli are able to affect hormone receptors could potentially provide a very useful general mechanism for adaptation in autonomous robots.

In this paper we present a simple model of such an epigenetic mechanism, in which the sensitivity of the receptors of simulated hormones is directly regulated by the hormone level. In our model we use varying simulated hormone concentrations, secreted as a function of the interactions of a robot with its environment, not only to “rewire” different neural or physiological functions as in e.g. [9], [10], [11], but also to regulate their receptors.

This regulation functions in a similar manner to a positive feedback loop. Higher concentrations of a hormone lead to upregulation (increased sensitivity) of a receptor and lower concentrations cause downregulation (reduced sensitivity). Here we investigate the adaptability of our epigenetic mechanism to a range of environmental challenges through allostatic processes [12] and demonstrate, in a series of experiments, how our system can lead initially identical robots (i.e. robots with the same initial internal parameter values and with the same speed potential) to develop different behavioral phenotypes well adapted to different environmental conditions, as a consequence of the changes brought about by and in the epigenetic mechanism in its interactions with the environment. In our model, no pre-defined behavioral repertoires or hand-crafted behaviors are used. Instead, all behaviors observed and discussed in the paper emerged as a result of both short-term (hormonal) and long-term (epigenetic) modulation of the perception and actuators of the robot.

The work presented here stems from, but expands very significantly, an earlier conference publication [13]. Here we provide a much deeper investigation by both adding new experimental conditions, analyzing the adaptive mechanisms of the robot more rigorously, and exploring the effect of experiential diversity on development in much greater detail. We refer the reader to [14] for a comprehensive look into the background of the approach.

The remainder of this paper is organized as follows. In section 2, we introduce the different aspects of our model. In section 3, we present the different experiments carried out to test our model and the results obtained. Finally, we draw some conclusions in section 4.

## 2 ROBOT MODEL

Our epigenetic robot architecture can be broken down into three systems, designed to build on each other incrementally:

- *Basic architecture*: the first system handles basic functions such as drives, motivational state, and movement (section 2.1).
- *Neuromodulatory system*: the second system modulates the basic architecture through the secretion and decay of four different simulated hormones (section 2.2).
- *Epigenetic mechanism*: the third system implements a simple epigenetic mechanism, which controls the sensitivity of the hormone receptors of the second system (section 2.3).

This architecture is designed incrementally: we can implement robots driven by the first system only (which we will call “Basic robot” in this paper), by a combination of the first and the second systems (called “Hormone-modulated robot”), or by the three interacting systems (called “Epigenetic robot”). The full (epigenetic) robot architecture, containing the three systems, is depicted in figure 1. This incremental design of the robot architecture permits a methodical, incremental testing and analysis of the functionalities and benefits (in terms of adaptation) that each subsequent addition brings to the robot. To this end, we tested each of the 3 robot models in the same (six) open-field environments posing different challenges, as we will explain in section 3.

### 2.1 Basic architecture

The basic behavior of the robot is driven by a combination of internal needs and external stimuli. The “basic architecture” is built around a motivational system that generates continuous behavior, running continuously in “action selection” loops (or time steps) of about 62 milliseconds (about 16 loops per second).

#### 2.1.1 Motivational system

Following Cañamero’s long-standing approach [15], our basic robot has a number of internal needs arising from a simulated physiology of homeostatically-controlled survival-related variables. Our robot has three simulated survival-related variables: energy, health and temperature (all with values between 0 and 100), which, to survive, it must maintain within a pre-defined range.

The variable *Energy* acts as a simulated battery that the robot can sense internally. This simulated battery has a maximum charge of 100 units (ideal value), a “fatal limit” (the robot “dies” if this is reached) of 0, and loses one unit of charge per second. The robot is able to recharge this battery by moving within a 2cm range of an “energy resource”—a pink ball.

The variable *Health*, a simulated measure of physical integrity, also has an ideal value of 100 units and a fatal limit of 0. Its actual value decreases with physical contact, which our robot detects with its 16 infrared sensors (IR), according to:

$$\Delta Health = \sum_i \begin{cases} -\frac{Ir_i}{IrMax} & \text{if } Ir_i > IrMax \times 0.95 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where  $Ir_i$  is the value of the IR  $i$ , and  $IrMax$  is the maximum sensor value, corresponding to an object touching the sensors. Hence, if all 16 IR detected an object in contact with the robot, then the robot would lose 16 health units per action loop. Like energy, this internal need can be recovered by consuming specific resources, in this case blue colored balls.

Finally, the (internal) variable *Temperature* depends on the current wheel speed of the robot and the simulated temperature of the external environmental, which we call *Climate*. It has an ideal value of 0 and a fatal limit of 100; therefore, the error of this variable is not by deficit but by excess, and coincides with the actual internal temperature. The temperature changes according to:

$$\Delta Temperature = \frac{|sp|}{10} \times Climate - (Temperature \times 0.05) \quad (2)$$

where  $|sp|$  is the absolute value of the wheel speed (measured in rotations per action loop) and the value 10 is a predetermined constant to regulate temperature gain.

Following a homeostatic model, the discrepancy between the actual and ideal values of an internal variable produces an error and a tendency to correct that error—a *Drive*. The robot has two explicit drives: a tendency to correct energy deficits, i.e. increase energy levels, ( $Drive_e$ ), which is satisfied by consuming “energy” (pink balls), and a tendency to repair (increase the level of) health ( $Drive_h$ ), which is satisfied by consuming “medicine” (blue balls). The need to decrease temperature is indirectly satisfied by reducing movement; this means that it can interfere with the satisfaction of the other drives, which involve movement of the robot. Therefore, these interactions need to be taken into account when calculating the intensity of the drives:

$$Drive_i = \frac{Deficit_i}{Error_t} \quad (3)$$

where  $Drive_i$  is the need to recover the homeostatic variable  $i$  ( $e$  or  $h$ ),  $Deficit_i$  is the deficit of the current homeostatic variable  $i$ , and  $Error_t$  is the temperature error.

The drives affect the robot’s perception of three different environmental cues using the scaling factors shown in table 1. In this paper, as the relationship between drive and cue is fixed, the robot’s perception of different cues can only be modulated by its internal state, allowing us to focus on the effects of the epigenetic model. However, in [16] we have integrated the epigenetic model into a developmental learning system permitting the robot to learn the affordances of different objects, i.e., to learn how it can interact with them, or the effect of its interactions.

In addition to its internal needs, the motivational system of the robot takes into account the external stimuli or environmental “cues” relevant to the satisfaction of its different needs (table 1) to decide on the appropriate course of action. Departing from previous models that used discrete motivations (e.g., [9], [15], [17], [18]), in

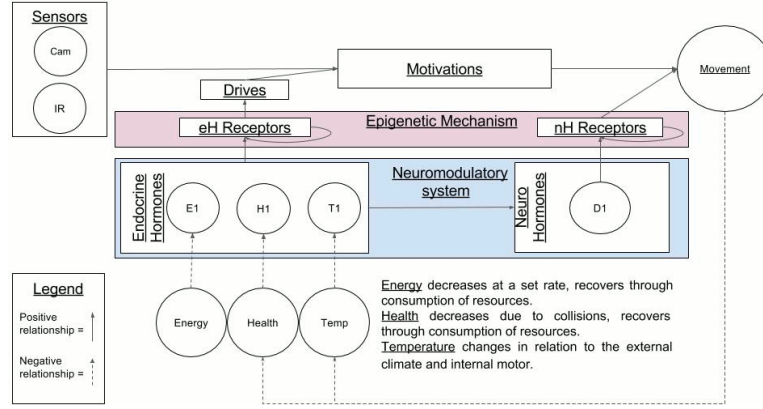


Fig. 1. Complete robot architecture with its three interacting “systems”: Basic architecture, Neuromodulatory system, and Epigenetic mechanism.

TABLE 1

Relationship between drives and environmental cues. The numbers shown are the scaling factors used in equation 4.

Drive	Environmental Cues		
	Energy Resource	Health Resource	Solid Object
Replenish energy	10	0	0
Repair health	0	10	-1

this architecture the robot has a global motivational state that will make it move at a particular speed and in a particular direction—one of eight directions around the robot, determined in body-coordinates. The motivational state of the robot (a vector) is calculated as follows:

$$MotState_j = \sum_{i \in \{e, h\}} \sum_c s_{i,c} (Drive_i \times Cue_{c,j}) - Avoid_j \quad (4)$$

where  $j$  is one of 8 directions around the body of the robot,  $s_{i,c}$  are the scaling factors in table 1,  $Cue_{c,j}$  is the magnitude of an environmental cue of type  $c$  in direction  $j$ , and  $Avoid_j$  is given by:

$$Avoid_j = \begin{cases} Ir_j & \text{if } Ir_{Max} - Ir_j < Drive_h \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

### 2.1.2 Behavior generation

The motivational state influences the generation of the behavior (motor action) to be executed. Unlike the majority of related motivational robot architectures previously developed in our group, which use discrete behaviors (e.g., [9], [15], [17], [18]), our robot uses a continuous sensory-motor mapping model, more in vein with models such as [19]. For example, rather than having an explicit “pushing” behavior, the robot may push an object when its “desire” to move forward outweighs its “desire” to avoid the object. Our model thus calculates the motion of the robot rather than selecting specific pre-defined behaviors. Motion is calculated independently for each wheel and the overall movement (speed and direction) results from the simultaneous movement of both wheels. The speed of the wheels, and therefore the observable behavior, is calculated by:

$$WheelSpeed_k = \sum_{j=0}^7 (MotState_j \times Set_{k,j}) \quad (6)$$

where  $WheelSpeed_k$  is the speed of the left or right wheel ( $k = 0$  or  $1$ ), and  $Set_{k,j}$  are vectors  $(-10, -10, -5, -3, 1, 3, 5, 10)$  if  $k = 0$  and  $(-10, 10, 5, 3, 1, -3, -5, -10)$  if  $k = 1$ , to translate the eight-directional motivational state vector into a coherent motor action.

As shown in equations 4 and 6, the robot’s movement is thus driven by both its internal state and external<sup>1</sup> stimuli. When the robot’s internal needs are satisfied, it will engage in an avoidance-like behavior and try to maintain a distance between itself and the objects perceived, stopping if it finds an empty<sup>2</sup> area. In contrast, when the robot needs to replenish health or energy but no resources are detected, it will move forward (avoiding obstacles) looking for them.

## 2.2 Neuromodulatory system

Our neuromodulatory system implements a hormonal system that modulates the basic motivational robot architecture. Although such simple motivational models can be effective behavior selection systems in simple environments, they tend to perform poorly in more complex and dynamic environments. A potential solution to this problem is found in biological systems in which motivations do not come directly from homeostatic errors. Rather, hormone secretion derived from homeostatic errors are behind the onset of motivation by providing a signal of the error [20] and the motivational value of environmental cues [21]. In our previous work, we have applied hormonal modulation to both the perception of external environmental cues and of internal errors (e.g. [9], [15], [18]) as well as to the internal elements of the architecture linking perception and actuation (e.g. [22], [23]). In this paper, (simulated) hormones secreted from

1. As we mentioned above and shown in table 1, the robot is able to detect three types of environmental cues that affect its motivational state and ultimately its behavior. For example, if the robot sees an energy resource and the intensity of its energy drive is 2, then its motivation to move towards the resource would be equal to  $20 \times Set_{k,j}$ .

2. Due to the range of the Koala robot’s IR sensors, an “empty area” roughly corresponds to a gap of 5cm between the robot’s IR sensors and any object.

TABLE 2  
The different hormones used in our architecture.

Hormone	Type	Secretion trigger	Regulates
E1	Eh	Energy error (deficit)	Energy recovery
H1	Eh	Health error (deficit)	Health recovery
T1	Eh	Internal temperature error (excess)	Speed
D1	Nh	Visual cues and Eh concentration	“Dominant / aggressive” behavior

(simulated) glands modulate different sensors and actuators of the robot. These hormones are secreted in relation to changes in the robot’s internal and external environment. Our model implements four hormones of two different types—endocrine hormones (Eh) and neurohormones (Nh)—as shown in table 2.

### 2.2.1 Endocrine hormones

Drawing on biological systems, our Eh hormones are implemented with the primary purpose of helping to maintain internal homeostasis. The Eh group consists of three hormones, each one associated with one of the three homeostatic variables. These hormones are secreted as a function of current homeostatic errors:

$$EhSecretion_i = \psi_i \times Error_i \quad (7)$$

where  $i$  is one of the three homeostatic variables, and  $\psi_i$  is a constant value to regulate secretion.

### 2.2.2 Neurohormones

We have implemented a single neurohormone (Nh), D1. This hormone facilitates what can be described as “dominant” or potentially “aggressive” behavior, and can be thought of as being somewhat similar to testosterone in mammals. This “dominant” behavior is achieved by the combination of two effects of Nh: (a) changing the wheel speed, as we will see in equation 11, and (b) reducing the tendency to avoid objects by modulating the perception of environmental cues that are associated with negative stimuli such as collisions, overheating or other actions which move the robot away from homeostatic balance (equation 10). This system shares some similarities with [9], where a hormone suppressed perceptual feedback from the robot’s bumpers (touch sensors), causing the robot to push competitors away as a side effect. However, in our model the Nh hormones have the effect of suppressing all negative stimuli. In addition, in our model the strength of suppression of negative stimuli is correlated with the concentration of hormones present in the system, as in [9], rather than with internal deficits, providing a simple form of “affective memory”. For example, a robot with a high D1 level that detects a desired resource will move directly towards it ignoring and pushing aside any obstacles in the way, disregarding the potential damage from collisions. In contrast, a robot in the same situation but with a low D1 level would instead move around obstacles to reach the

desired location. Another example is the ability of the Nh to modulate the robot’s speed and suppressing the drive to slow down due to overheating.

The amount of secretion of the Nh is determined by the intensity of all the internal drives, as given by:

$$NhSecretion = \sum_i Drive_i \quad (8)$$

### 2.2.3 Hormone Receptors

Once secreted, both types of hormone, Eh and Nh, enter the virtual bloodstream that surrounds the architecture and persist for a random number of action selection loops (within a fixed range) before decay of that particular secretion occurs. Hormone receptors can detect the different concentration of these two types of hormones. These receptors are then able to signal for the appropriate modulation to internal or environmental change. This modulation is achieved by modifying equation 3 to:

$$Drive_i = \frac{Sens_i \times Eh_i}{Sens_t \times Eh_t} \quad (9)$$

where  $t$  is temperature,  $i \neq t$  is one of the other two variables, and  $Sens_i$  is the sensitivity of the receptor to the level of the corresponding hormone. In the epigenetic robot (seen section 2.3),  $Sens_i$  is modulated by the epigenetic mechanism, but in the hormone-modulated robot,  $Sens_i$  is a constant = 1.

The suppression of negative stimuli by the Nh is achieved by modifying equations 5 and 6 respectively to:

$$Avoid_j = \begin{cases} \frac{Ir_j}{Nh} & \text{if } IrMax - Ir_j < Drive_h \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

$$WheelSpeed_k = \sum_{j=0}^7 (MotState_j \times Set_{k,j}) Sens_{Nh} \times Nh \quad (11)$$

## 2.3 Epigenetic system

The final system of the model introduces an epigenetic adaptation mechanism into the architecture. Taking inspiration from biological studies (see [5], [6], [8] for an overview) in which hormone concentrations are shown to correlate with the triggering of varying epigenetic mechanisms that influence the expression of hormone receptors, we use simulated hormones in our model to regulate the robot’s hormone receptors. High concentration of a hormone leads to the upregulation of the receptors (i.e., an increased number of receptors) and lower concentrations cause downregulation (i.e., a lowered number of receptors). As we show in this paper, changes in regulation of receptors lead to the emergence of persistent behaviors relevant to the robot’s current environment. These behaviors can of course still be modulated by variations in hormone concentration levels.

In this work we have not tried to implement an explicit model of the epigenetic process (i.e., an explicit model of DNA methylation), but rather we have modeled the outcome of the process by updating the previous value of the  $Sens_i$  in equation 9 as follows:

$$Sens_i(t+1) = Sens_i(t) \frac{Eh_i}{\sigma} \quad (12)$$

where  $\sigma$  is a predetermined constant value that regulates the speed of the epigenetic process,  $Sens_i(t)$  is the sensitivity of the receptor for hormone  $i$  at time step  $t$ , and  $Eh_i$  is the hormone concentration; hence, the hormone influences the upward or downward regulation of its own receptors.

### 3 EXPERIMENTS

To assess the ability of the epigenetic robot to adapt to different environmental challenges, we tested it in six different environments, where we also tested the basic and neuromodulated robots for comparison. The purpose of these experiments was to investigate the adaptability of our epigenetic mechanism to a range of environmental challenges through allostatic processes [12], and more specifically how our system can allow initially identical robots (with the same initial internal parameter values and speed potential) to develop different behavioral phenotypes adapted to different environments as a consequence of the changes brought about by and in the epigenetic mechanism in its interactions with the environment. All the adaptations took place as a result of such dynamics of interactions, without the need of programming or tuning the robot to operate in a specific environment.

Adaptation was assessed in terms of how well the internal physiological variables were maintained, quantitatively measured in terms of “comfort level” [9] (figures 4, 5, 6, 7, 9 and 10), which we calculated as the normalized mean of the robot’s *Energy*, *Health* and  $100 - \text{Temperature}$ .

These experiments were carried out using the medium-sized wheeled Koala-II robot<sup>3</sup>. In addition, a webcam was added to the robot to allow for simple color-based vision. Computations were performed on a laptop fitted on top of the robot. Each “action selection” loop lasted 62.5 milliseconds, equivalent to 16 loops per second.

The experiments were performed inside a  $2m \times 2m$  arena with wooden borders, shown in figure 2. Each of our first four experiments consisted of 10 runs for each robot architecture, with a duration of 10,000 steps (action selection loops) each, or approximately 10 minutes at a rate of 16 steps per second. The last two experiments involved 15 runs of 10,000 steps for each architecture. The addition of 5 extra runs for last two experiments was due to increased environmental variation under these two conditions (sections 3.5 and 3.6). In each experiment, the robots had access to at least one item of each resource, with the exact number and environmental challenges changing significantly across experimental sets. If at any point one of the robot’s homeostatic needs passed its fatal limit, it “died” and the run would end.

Table 3 provides an overview of each experiment, table 4 shows the physiological changes that the epigenetic robot underwent in terms of receptor regulation in the six experiments, and figure 3 shows the evolution of the



Fig. 2. The Koala robot used in our experiments. Additionally, the light movable objects from experiment 2 and a moving “prey” resource from experiment 3 can be seen at the top and left of the figure respectively.

sensitivities of the hormone receptors over the course of the experiments and observed behaviors.

#### 3.1 Results of experiment 1: basic environment

In this simple environment all three robot architectures were able to maintain homeostatic balance (see figure 4). However, while the three models were similar in terms of homeostatic maintenance, they varied in their behavior.

The basic architecture and hormone-modulated robots behaved as expected: homeostatic errors would cause modulation of the wheels (cf. equations 6 and 11) and motivate the robots to move towards and consume the appropriate resources. In contrast, the epigenetic robot stayed close to an energy resource (all located in a corner) at all times even when not consuming it, either stopping nearby or tightly circling it. This behavior occurred due to a downregulation of the H1 and T1 hormone receptors shown in table 4 and figure 3. Limited presence of the H1 and T1 hormone due to the simplicity of the environment (few obstacles and limited required movement) resulted in this change in expression. As the energy decreased at a set rate, the E1 receptor remained constant. By staying stationary, this robot was able to almost entirely eliminate the other drives. As can be seen in figure 4, while results in terms of adaptation of all robots (as measured by comfort levels) are close, the epigenetic robot maintained slightly higher overall values particularly in the latter stages, supporting idea that in this case lack of movement was adaptive.

#### 3.2 Results of experiment 2: movable objects

In this scenario, the non-epigenetic robots performed poorly (see figure 5): the basic architecture robot died in all runs and, while the hormone-modulated model performed better, with only 3 early deaths overall, its performance was still poor. The inability of these architectures to adapt to the harsher environment stemmed from two causes. Firstly, they often failed to maintain sufficient health to push past obstacles to reach the resources (i.e., health deficits led to overriding collision avoidance). Secondly, they were unable to adapt to the longer time needed to find a resource due to the need to navigate obstacles. These issues resulted in both low energy and poor health, whereby robots “died” due to either running out of energy, or from impact of pushing past objects.

3. <http://www.k-team.com/mobile-robotics-products/old-products/koala>



TABLE 3  
Overview of the different experiments and the challenges they posed to the robot .

Experiment	Challenge
1) Basic Environment	An open environment with no significant challenges, used as a baseline for comparison with more complex environments. Two items of each resource were present at all times.
2) Light Objects	The environment contains light objects that the robot needs to push past in order to reach resources. As pushing results in a loss of health, this creates a trade-off for the robot between maintaining its energy or its health.
3) Moving Resources	The energy resources move in a constant pattern at a set speed, attempting to simulate a simple prey. As they move at the same speed as the robot, the robot cannot simply chase after them to feed.
4) Dynamic Climate	The temperature of the environment changes through the experiment in a day-night cycle, with hotter temperatures during the “day” and slowly cooling down at “night”. This challenges the robot to satisfy drives during cooler “night” periods so they can reduce movement during the day and not overheat.
5) Uneven Resources	Three sets of 5 runs with uneven resource distribution in each set of runs. Each set placed stress on a specific homeostatic variables (e.g., one set of runs would be in an environment with a reduced number of energy resources). This environment challenges the robot to adapt to prioritize less abundant/harder-to-maintain needs.
6) Temporal Dynamics	The resources are only available for limited periods of time. This forces the robot to take opportunities to recover deficits when the resources appear. For the first five sets of runs, the energy source would be present for 30 seconds per minute. In the second five sets of runs the duration was reduced to 20 seconds. In the final five sets of runs the resource was only accessible for 10 seconds of every minute. As only one type of resource is affected the robot must prioritize this one over the others.

TABLE 4  
Receptor regulation of the epigenetic robot. An unregulated receptor would have a value of 1, upregulated  $>1$ , downregulated  $<1$ .

Experiment	E1 Receptor	H1 Receptor	T1 Receptor	D1 Receptor
1) Basic Environment	1.01	0.69	0.86	0.41
2) Light objects	1.21	1.34	1.02	1.29
3) Moving Objects	1.43	1.13	1.56	1.33
4) Dynamic climate	1.31	1.03	1.78	0.91
5) Uneven Resources set 1	0.89	1.14	0.91	0.86
set 2	1.25	0.72	1.07	0.82
set 3	1.06	1.01	1.22	0.80
6) Temporal Dynamics	1.54	0.86	0.94	1.21

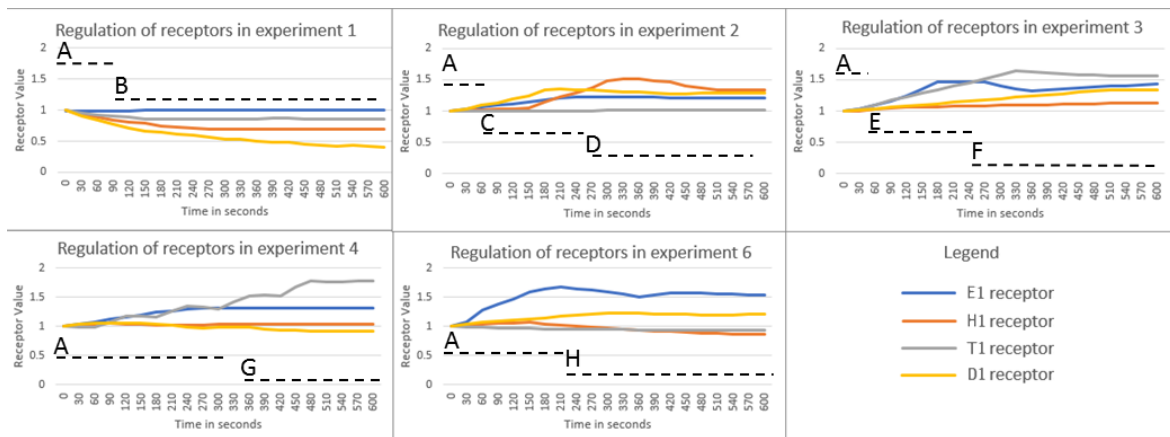


Fig. 3. Graphs showing the changes in sensitivities of the hormone receptors in the epigenetic robot over time. Letters indicate the occurrence of observed robot behaviors that will be discussed in the experiments (sections 3.1 to 3.6): A) Simple reactive behavior. B) Due to downregulation of D1, the robot is attracted to only one aspect of the environment using a guarding behavior. C) Reduced object avoidance due to upregulation of D1, robot pushes past obstacles. D) Pushing behavior refined, as a result of upregulation of H1. The robot now repairs before attempting to push. E) Robot chases after resources, slowly gaining speed as motivation to reach them increases until it reaches them or overheats and stops. F) Upregulation of T1 means the robot using an ambush strategy, staying still until until motivation to move to a resource increases (due to close proximity) at which point upregulation of E1/D1 ensures the robot moves at full speed. G) Upregulation of T1 results in a “hibernation” behavior, exploiting favorable weather condition to recover homeostatic deficits before returning to a dormant state. H) Rapid upregulation of E1 leads the robot to a state of extreme sensitivity to opportunities to recover energy. This heightened sensitivity will override any other needs, allowing the robot to take advantage of the limited windows of opportunity to recover energy. Upregulation of D1 means that the robot will move at higher speed towards the resources.

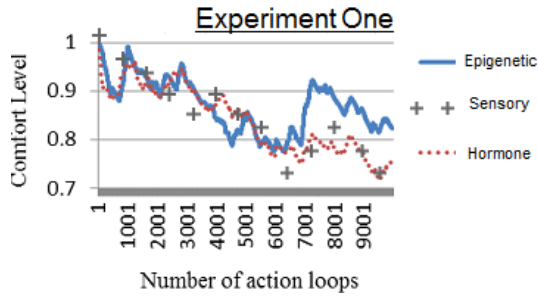


Fig. 4. The results of the first experiment (basic environment).

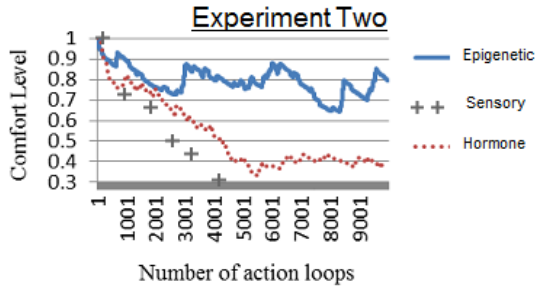


Fig. 5. The results of the second experiment (light objects).

In comparison, the epigenetic robot was able to adapt to the new environment due to a chain reaction of internal changes (see table 4). Firstly, perception of resources that were visible but inaccessible led to persistent high levels of the D1 hormones and an upregulation of the D1 receptor. The increase in D1 sensitivity made this robot more likely to push past obstacles rather than going around. This behavior resulted in increased health deficits, and subsequently increased concentrations of the H1 hormone and the upregulation of the H1 receptor. This increased the motivation to replenish health after pushing past an object, and hence led to maintenance of high levels of health. These epigenetic changes ensured that this robot could safely push past objects when needed.

We also observed differences in the strength with which the robots “pushed” objects. The way in which the robots changed the force of their push was by driving faster or slower into an object<sup>4</sup>. The epigenetic robot developed a more forceful or intense “pushing” behavior due to upregulation of D1, which on average resulted in 12% less health loss than the slower pushes used by the other robots. Forceful pushes also increased the likelihood of the epigenetic robot to fully complete the pushing behavior, pushing completely past an object rather than giving up and moving away part way through. In this experiment, the average speed of the epigenetic robot while trying to push was 230mm/s, in comparison to 150mm/s for the hormone-modulated and 60mm/s for the basic architecture robots.

4. The speed of the push of the robots depended on the intensity of the temperature drive (T1), which decreased speed, and D1 hormone concentration, which increased speed (see equations 9 and 11 respectively).

### 3.3 Results of experiment 3: moving resources

In this environment the resources moved in regular patterns at the default speed of the robots, simulating a simple prey. Due to their similar speeds, the robots could not simply chase a resource to reach it and needed to develop a “hunting” behavior to avoid drawn-out “chases” that could make them overheat critically. As seen in figure 6, the epigenetic robot’s behavior proved most suitable. This robot developed an “ambush-like” strategy, where it would remain sedentary until a needed resource passed nearby, at which point it would “pounce”, giving chase at full speed. This behavior emerged due the following mechanisms: Overheating from initial “chases” resulted in significant upregulation of the T1 receptor increasing sensitivity to its temperature and subsequently in suppressed movement, which led to a low internal temperature. The robot was able to overcome this suppression, and thus move, either with very high E1/H1 concentrations or if a resource came within close distance<sup>5</sup>. Due to the finely balanced internal parameters and the robot’s low internal temperature, once the motivation to move developed the robot was able to move at full speed for short bursts. Due to the speed of the epigenetic robot when it gave chase, the resource was often pinned against a wall, making it easier for the robot to feed on it; this also reduced the length of the “chases”, which lasted only around 4 seconds and were successful 87% of the time.

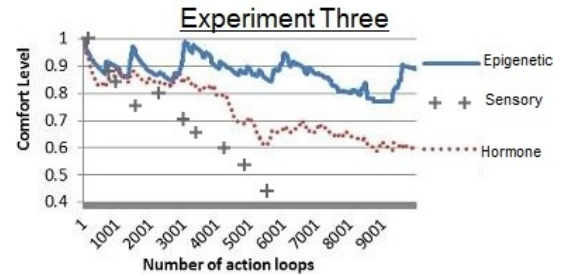


Fig. 6. The results of the third experiment (moving resources).

The hormone-modulated model also had a high success rate of 72% but showed longer chases of around 14 seconds on average. Compared with the epigenetic robot, the hormone-modulated robot engaged in long chases, often catching the resource when it turned in a corner. During these chases this robot’s speed would increase the closer it got to the resource (more intense stimulus) and the longer the chase went on (increased D1 concentration). Failure here was always due to the fact that the robot overheated before it reached the resource and needed to suspend or slow down movement in order to reduce its temperature.

Finally, the basic architecture robot only had a success rate of 13%, with an average chase taking 12 seconds. This robot would follow a needed resource; however, due to the lack of the D1 hormone it could rarely catch up with a resource. Unless the robot happened to be near one of the end points where the resource stopped, it would abandon the chase due to overheating. The inability to catch resources

5. By the end of a run, assuming base levels of the E1 hormone, this distance was around 12cm, with higher concentrations furthering it.

meant that this robot constantly had critically low levels of the survival-related internal variables and in all cases died before the end of the experiment. Due to its low success rate, the basic architecture robot also spent a larger amount of time engaging in chases—67% of its time in comparison to 28% for the hormone-modulated robot and just 7% for the epigenetic robot.

### 3.4 Results of experiment 4: dynamic climate cycle

As can be seen in figure 7, the basic architecture robot performed poorly in this environment as it died before the end of the second climate cycle, with causes of death evenly split between lack of energy and overheating. Simply put, this robot did not adapt to take advantage of the colder period, and still had energy or health deficits during the hot period. This resulted in a loose-loose situation: either moving to replenish energy or health and dying of overheating, or staying still and running out of energy.

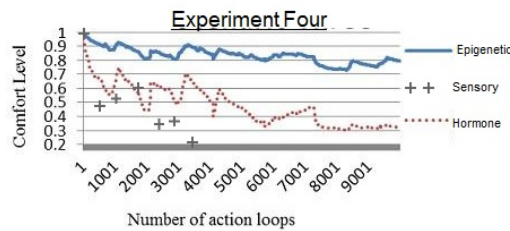


Fig. 7. The results of the fourth experiment (dynamic climate).

The hormone-modulated robot fared better. While it was still unable to fully take advantage of the cold periods, a slow-paced moving behavior emerged during the hot periods due to a combination of high T1 and medium levels of E1 and H1. However, even with this behavior, three deaths were still recorded due to energy loss.

Finally, the epigenetic robot developed two contrasting behaviors in order to survive the periods of high ambient temperature. In some (seven) runs, it developed a “hibernation” behavior: the robot became highly attracted to the different resources during the cold period and thus would fully replenish any deficits before laying dormant during the hot periods. This “hibernation” behavior emerged due to upregulation of the three Eh receptors (table 4). In particular, upregulation of the T1 receptor (due to the heat during the day) resulted in the robot completely suppressing movement during these periods. As the robot could not move during the day, in the early stages deficits in health or energy could not be corrected during this period, resulting in rising concentrations and therefore upregulation of their respective receptors. When the period of heat passed, this robot would immediately seek to correct these deficits regardless of how small they were. The second, alternative behavior that the robot developed (in 3 runs) was to simply stay near the energy source at all times, except when the occasional need to repair arose; this behavior permitted the robot to continue consuming energy during the hot temperature climate period, with only very limited movement needed. This behavior has a similar basis to the “hibernation” behavior observed, but was linked

to upregulation of T1 and E1 receptors, i.e., there was no upregulation of the H1 receptor. The behavior adopted (hibernation or staying near) depended on the robot’s early life experiences, and specifically on the amount of health loss that occurred during the first few cycles. If a robot completed these cycles with minimal health loss, then the second behavioral phenotype was adopted. If on the other hand a robot had health deficits (from early collisions), it adopted the “hibernation” behavior. Interestingly, it appeared that after the initial cycles and the adoption of one of these two behaviors, future health loss (or lack of health loss) did not result in the robot changing phenotype.

As shown in figure 8, the epigenetic robot explored a large portion of the environment while maintaining a low temperature. In contrast, the two other robots moved primarily between the two resources located in the top left and bottom right corners but had high internal temperatures. For the hormone-modulated robot, these high temperature periods were normally a result of running low on energy and needing to move quickly to replenish. For the basic architecture robot, however, these periods simply represented a lack of adaptation to the increased ambient temperature, as it continued their normal behavior even in the hot conditions.

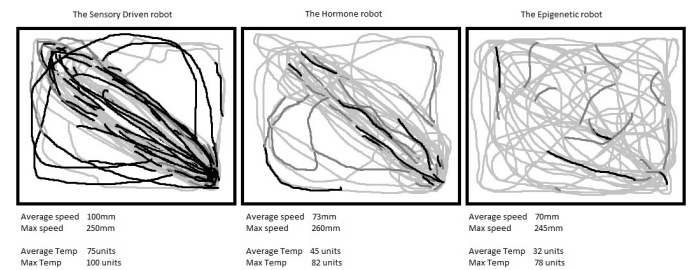


Fig. 8. Movement maps of the three different robots. The color of the lines represents their internal temperature: light gray a temperature between 0–50 units, dark gray 51–75 units, and black 76–100 units.

### 3.5 Results of experiment 5: uneven resources

Results are shown in figure 9. The basic architecture model was in all cases able to operate at a low but passable level. While there were four deaths spread across the experiments and cases of low comfort levels, this was the only time, excluding the first experiment (section 3.1), when a majority of the basic architecture robots survived. Behaviorally, there is little to report, as it behaved almost exactly the same as in experiment one.

The hormone-modulated model performed better but still had a hard time dealing with uneven resources. The overload of stimulation from the excess resources meant that the homeostatic deficit correlating to the rare resources was on average 32% greater. On occasions when the rare resource was hidden behind another resource, the robot would stop searching for it, as it became distracted by other needs.

In contrast, the epigenetic model was able to successfully adapt. Early homeostatic deficits led to the rapid secretion of hormones triggering early epigenetic changes and



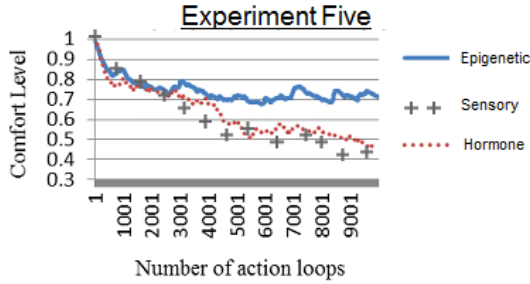


Fig. 9. The results of the fifth experiment (uneven resources).

increasing sensitivity of the associated hormone receptors. This in turn made the robot more sensitive to these deficits, filtering out the distraction and noise of the other resources. When both stimuli were present, the deficit of the overabundant resource would need to be around 42% greater in order to attract the robot's attention. Further, in the four cases when the rare resource was hidden, the deficit related to the common resource would need to drop to between 60–80% before the search for the rare resource was interrupted.

Finally, the set of experiments with the high level of temperature (set 3) did not yield considerable differences in terms of homeostatic comfort levels between the three models. While the epigenetic model performed better in the experiments, the biggest difference was qualitative in terms of the behaviors that emerged from the models. The epigenetic model developed a “stalking-like” approach due to upregulation of the T1 and E1 receptors. This behavior emerged in a similar manner to the “pounce” behavior seen in the third experiment (section 3.3) except that the degree of upregulation was smaller (see table 4). This meant that movement was not completely suppressed and the robot displayed a very slow exploration behavior until a resource was detected, at which point it moved quickly. In contrast, the hormone-modulated model would maintain a slow constant speed due to equal levels of the E1 and T1 and move towards the desired resource. While both behaviors worked, the approach of the epigenetic robot generated less overall heat, and was therefore slightly preferable for this environment.

### 3.6 Results of experiment 6: temporal dynamics

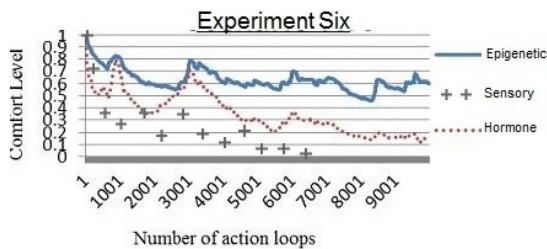


Fig. 10. The results of the sixth experiment (temporal dynamics).

As can be seen in figure 10, the basic architecture model was unable to deal with the temporal dynamics, never managing to survive a full run. In the first five sets of runs, when the resource was available for 30 seconds, the

robot was able to survive around eight minutes on average using the same behavior as in experiment one (section 3.1). However, as the period for which the temporal resource was available decreased, the robot increasingly struggled and by the final five sets of runs all agents had died before the 2-minute mark. The main problems for this robot were its inability to adapt to “rare” conditions, as in experiment five (section 3.5), and the inability to adapt its speed in order to move promptly towards the temporal resource. Overall, when a temporal resource was available, the basic architecture model managed to reach it only 17% of the time on average. In contrast, the average success rate was 84% for the epigenetic and 62% for the hormone-modulated models.

Comparing the epigenetic and hormone-modulated models proved more interesting. Both robot types performed at a similar level during the first five runs with the 30-second window of opportunity. While the epigenetic robot moved more promptly to resources when they appeared due some upregulation of the E1 receptor, neither robot was ever in any real danger of missing an opportunity. For the epigenetic robot, upregulation of the E1 receptor occurred as a result of increased energy deficits and subsequent E1 hormone concentrations, due to the fact that the resource was not always available when needed. However, as the window of opportunity shrank, the differences between the two models became apparent. Since the point where the resource would appear next was unknown to the robots, it was inevitable that both architectures would now miss some opportunities to replenish, although the epigenetic robot was generally quicker to move to the energy resource due to heightened sensitivity as a result of the upregulation of the E1 receptor. Finally, due to missed opportunities to fully recover deficits, both robots often had high levels of the D1 hormone. This in turn resulted in increased aggressive-like behavior and the occurrences of collision in later runs, subsequently increasing the need for the repair resources. In multiple cases this lead to similar or greater deficits in health in comparison to energy, and therefore to similar or greater hormone concentrations of H1 in comparison to E1. As health deficits could be, and were readily recovered in the epigenetic robot, they did not lead to any significant upregulation of the H1 receptor. However, the health deficits did cause the hormone-modulated robot to sometimes go to the readily available repair resource during the limited periods when the energy source was present and seen. This was due to the similar effects that both hormone concentrations had on this robot. In contrast, the epigenetic model, due to the upregulation and heightened sensitivity of the E1 receptor had effectively adapted to the rarity of the resource. Even with health deficits present, it only missed the opportunity to replenish energy once, when its health levels were critical (both under 10% and at least 63% less than energy). In total, seven of the hormone-modulated robot runs ended prematurely compared to a single death in the epigenetic model.

## 4 CONCLUSION

Drawing inspiration from development in biological organisms, an epigenetic process in which environmental

stimuli are able to affect hormone receptors could potentially provide a very useful general mechanism for adaptation in autonomous robots. We have implemented a simple model of such an epigenetic mechanism, in which the sensitivity of hormone receptors is directly regulated by the hormone level, and our results show that it successfully leads to rapid adaptation to different environmental challenges. Our experiments show that this epigenetic mechanism significantly and consistently improves the robot's adaptability in comparison to a basic homeostatically-controlled motivated architecture, and an improved version of this basic architecture that uses (simulated) hormonal modulation. In all our experiments, the epigenetic robot proved better at maintaining homeostasis than the other models. In addition, the variance between the runs in experimental sets of the epigenetic robot was considerably lower, indicating more consistent performance. This improved adaptation ability resulted from changes in the regulation of hormone receptors that in turn changed the robot's sensitivity to the variations of the simulated hormones, which signaled to the robot both its internal state and its surrounding environment. These changes to receptors led to the development of different tolerances and sensitivities to the internal and external conditions of the robot (for example, in an environment with plentiful energy supply the robot would tolerate larger energy deficits) that resulted in suitable adaptation of the robot to different environments, shown in both appropriate (and measurable) regulation of its internal survival-related variables and the emergence of successful functional behaviors and tactics such as pushing or hibernation. In each case, at the start of the experiments identical robots (i.e., robots with similar internal parameter values) were used, and the described adaptive qualities and behaviors were achieved "in real time" and without the need to fine-tune the architecture for the experiments.

The versatile and robust adaptation capabilities shown by our epigenetic architecture in the six environments were it was tested, which posed very different challenges for the robot, supports our view that an epigenetic mechanism can provide a very useful general mechanism for adaptation in autonomous robots. This is corroborated by the application of our epigenetic model to other contexts, such as human-robot interaction [16], [24], formation of group structures [25], and coupled to a neural network to modulate both the learning and behavior of a robot [16].

## ACKNOWLEDGMENTS

John Lones was funded a PhD studentship from the University of Hertfordshire. Partial support was also provided by the FP7 project ALIZ-E to Lola Cañamero.

## REFERENCES

- [1] J. L. Krichmar, "A biologically inspired action selection algorithm based on principles of neuromodulation," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2012, pp. 1–8.
- [2] G. A. Forehand and G. Von Haller, "Environmental variation in studies of organizational behavior," *Psychological Bulletin*, vol. 62, no. 6, p. 361, 1964.
- [3] H. A. Simon, *The Sciences of the Artificial*. Cambridge, MA/London: MIT Press, 1969.
- [4] V. Braitenberg, *Vehicles: Experiments in Synthetic Psychology*. MIT Press, 1984.
- [5] A. L. Fowden and A. J. Forhead, "Hormones as epigenetic signals in developmental programming," *Experimental Physiology*, vol. 94, no. 6, pp. 607–625, 2009.
- [6] D. Crews, "Epigenetics, brain, behavior, and the environment," *Hormones (Athens)*, vol. 9, no. 1, pp. 41–50, 2010.
- [7] F. A. Champagne, "Epigenetic influence of social experiences across the lifespan," *Dev Psychobiol*, vol. 52, no. 4, pp. 299–311, 2010.
- [8] X. Zhang and S.-M. Ho, "Epigenetics meets endocrinology," *Journal of molecular endocrinology*, vol. 46, no. 1, pp. R11–R32, 2011.
- [9] L. Cañamero and O. Avila-García, "A bottom-up investigation of emotional modulation in competitive scenarios," in *Affective Computing and Intelligent Interaction, Second International Conference, ACII 2007, Lisbon, Portugal, September 12–14, 2007, Proceedings, 2007*, pp. 398–409.
- [10] J. Timmis, L. Murray, and M. Neal, "A neural-endocrine architecture for foraging in swarm robotic systems," in *NICSO 2010, ser. Studies in Computational Intelligence*, 2010, vol. 284, pp. 319–330.
- [11] J. L. Krichmar, "A neurobotic platform to test the influence of neuromodulatory signaling on anxious and curious behavior," *Frontiers in Neurobotics*, vol. 7, 2013.
- [12] J. Schulkin, Ed., *Allostasis, Homeostasis, and the Costs of Physiological Adaptation*. Cambridge University Press, 2004.
- [13] J. Lones and L. Cañamero, "Epigenetic adaptation through hormone modulation in autonomous robots," in *IEEE 3rd Joint Intl Conf on Development and Learning and Epigenetic Robotics (ICDL)*. IEEE, 2013.
- [14] J. Lones, "Hormonal modulation of developmental plasticity in an epigenetic robot," Ph.D. dissertation, University of Hertfordshire, U.K., 2017.
- [15] D. Cañamero, "Modeling motivations and emotions as a basis for intelligent behavior," in *Proceedings of the First International Symposium on Autonomous Agents (Agents'97)*. ACM Press, 1997, pp. 148–155.
- [16] J. Lones, M. Lewis, and L. Cañamero, "From sensorimotor experiences to cognitive development: Investigating the influence of experiential diversity on the development of an epigenetic robot," *Frontiers in Robotics and AI*, vol. 3, no. 44, 2016.
- [17] I. Cos, L. Cañamero, G. M. Hayes, and A. Gillies, "Hedonic value: enhancing adaptation for motivated agents," *Adaptive Behavior*, pp. 465–483, 2013.
- [18] M. Lewis and L. Cañamero, "Hedonic quality or reward? a study of basic pleasure in homeostasis and decision making of a motivated autonomous robot," *Adaptive Behavior*, vol. 24, pp. 267–291, 2016.
- [19] A. J. Blanchard and L. Cañamero, "Developing affect-modulated behaviors: Stability, exploration, exploitation or imitation?" in *Proceedings of the Sixth International Workshop on Epigenetic Robotics*, ser. Lund University Cognitive Studies, vol. 128. Paris, France: Lund University, 2006, pp. 17–24.
- [20] S. Malik, F. McGlone, D. Bedrossian, and A. Dagher, "Ghrelin modulates brain activity in areas that control appetitive behavior," *Cell Metabolism*, vol. 5, no. 0, pp. 400–409, 2008.
- [21] N. H. Frijda, *The Emotions*, ser. Studies in Emotion and Social Interaction. Cambridge University Press, 1986.
- [22] R. L. French and L. Cañamero, "Introducing neuromodulation to a braitenberg vehicle," in *Procs 2005 IEEE Int Conf on Robotics and Automation*. IEEE, 2005, pp. 4188–4193.
- [23] K. Parussel and L. Cañamero, "Biasing neural networks towards exploration or exploitation using neuromodulation," in *Proc. 17th International Conference on Artificial Neural Networks (ICANN 2007)*, Part II, ser. LNCS, J. M. de Sá, L. A. Alexandre, W. Duch, and D. Mandic, Eds., vol. 4669. Porto, Portugal: Springer Berlin Heidelberg, 2007, pp. 889–898.
- [24] J. Lones, M. Lewis, and L. Cañamero, "Hormonal modulation of development and behaviour permits a robot to adapt to novel interactions," in *ALIFE 14: The Fourteenth Conference on the Synthesis and Simulation of Living Systems*, vol. 14, 2014, pp. 184–191.
- [25] —, "Hormonal modulation of interaction between autonomous agents," in *4th International Conference on Development and Learning and on Epigenetic Robotics*. IEEE, 2014, pp. 402–407.



**John Lones** is currently a Research Fellow in Bio-Inspired Autonomous Robots at the University of Leeds and part of the interdisciplinary ESPC funded Self Repairing Cities project. Additionally he is also visiting lecturer at the University of Hertfordshire. He holds a BSc degree in Economics, a MSc in Computer Science and a PhD in Computer Science (Artificial Intelligence) from the University of Hertfordshire (2017). He carried out his PhD at the Embodied Emotion, Cognition and (Inter-)Action Lab. His research interests include Embodied Artificial Intelligence, Epigenetics, Emergent Development and Behavior and Bio-Inspired Deep-Learning. Website: [JohnLones.com](http://JohnLones.com)



**Matthew Lewis** is an Early Career Research Fellow in Autonomous Robots as Embodied Models of Mental Disorders at the University of Hertfordshire, working in the Embodied Emotion, Cognition and (Inter-)Action Lab. He holds a BA in Mathematics from Cambridge University, and an MSc and PhD in Mathematics from Imperial College, London. Before returning to academia, he worked in the private sector as a Software Engineer on embedded systems. His research interests include Embodied Artificial Intelligence, robot models of mental disorders, and artificial emotions in robots.



**Lola Cañamero** is Reader in Adaptive Systems and Head of the Embodied Emotion, Cognition and (Inter-)Action Lab in the School of Computer Science at the University of Hertfordshire in the UK, which she joined as faculty in 2001. She holds undergraduate (Licenciatura) and post-graduate degrees in Philosophy from the Complutense University of Madrid and a PhD in Computer Science (Artificial Intelligence) from the University of Paris-XI. She turned to Embodied AI and robotics as a postdoctoral fellow in the groups of Rodney Brooks at MIT and of Luc Steels at the Free University of Brussels. Since 1995, her research has investigated the interactions between motivation, emotion and embodied cognition and action from the perspectives of adaptation, development and evolution, using autonomous and social robots and artificial life simulations. Some of this research has been carried out as part of interdisciplinary projects where she has played Principal Investigator and coordinating roles, such as the EU-funded HUMAINE, FEELIX-GROWING and ALIZ-E, or currently the UH-funded Autonomous Robots as Embodied Models of Mental Disorders. She is author of over 140 peer-reviewed publications in these topics, and played a pioneering role in nurturing the emotion modeling community through activities such as workshop and conference organization, a mailing list, and editorial activities. Website: <http://www.emotion-modeling.info>.