



Using habituation as a simple and fundamental learning mechanism in an embodied artificial agent

Tristan Gillard, Jérémy Fix, Alain Dutech

► To cite this version:

Tristan Gillard, Jérémy Fix, Alain Dutech. Using habituation as a simple and fundamental learning mechanism in an embodied artificial agent. *Adaptive Behavior*, In press, 31 (4), pp.299-321. 10.1177/10597123221116183 . hal-04031614

HAL Id: hal-04031614

<https://hal.science/hal-04031614v1>

Submitted on 16 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Using habituation as a simple and fundamental learning mechanism in an embodied artificial agent

Journal Title
XX(X):1–22
©The Author(s) 2021
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/

SAGE

Tristan Gillard^{1,2}, Jérémy Fix^{1,3} and Alain Dutech^{1,4}

Abstract

Habituation, a non-associative learning widely observed across phylogeny, is fundamental for adaptation and, thus, survival of living organisms. This paper investigates the main characteristics of habituation in order to present three new computational models inspired by habituation. We develop these models as part of the Iterant Deformable Sensorimotor Medium (IDSM), a recently developed abstract model of behavior formation. The characteristics of these models are studied and analyzed. Our long term objective is to research new unsupervised learning mechanisms for artificial learning agents.

Keywords

Non-associative learning, habituation, artificial agent, adaptive behavior acquisition, embodiment

1 Introduction

Habituation is a widespread form of learning in living organisms (Rankin et al. 2009). Habituation can be approximated as the decline in response to a repeated stimulus, and we describe in more details what habituation is in section 2. In addition to mammals, habituation has also been observed in animals without a central nervous system and even in unicellular organisms, e.g. the blob *Physarum polycephalum* (Vogel and Dussutour 2016) and other protozoa (Eisenstein et al. 2001). This suggests the hypothesis that habituation is a fundamental mechanism for any living organism, as one of the ways to adapt to changes in its environment. Indeed, the Rescorla-Wagner model of classical conditioning, a more complex learning process, is based in part on the existence of a habituation mechanism (Hall and Rodríguez 2020; Vogel et al. 2019).

As such, habituation could be used as an inspiration source to incorporate new adaptation mechanisms into artificial agents or robots facing real world challenges provided we have access to computational models of habituation. The objective of this work is thus to propose computational models inspired by habituation and analyse their properties. In a future work, we will explore how these models can help artificial agents acquire and develop new behaviors.

Habituation is a behavioral property which has been studied for quite a long time (Thompson 2009). The work of Rescorla-Wagner led to the Sometimes Opponent Processes model (SOP) (Wagner 1981). In this conceptual model, habituation is the result of some elements representing the stimulus being in “refractory state” (state A2 of the SOP model) for some time after the presentation of a stimulus. Thus, when the stimulus is presented again shortly after, these elements can no longer contribute to the response behavior, lessening its intensity. Other mathematical models of habituation using differential equations have been explored by (Stanley 1976) and refined later by (Wang 1993) in order to incorporate long-term effects.

More operational models have been proposed, mostly using neuron like formalism. A first approach consists in incorporating a sensor memory which builds up with the repeated presentation of a stimulus and that ultimately modulates the expressiveness of a motor response. This approach has for example been explored in (Marsland et al. 1999). In their paper, the authors consider both a simple linear integrator to modulate the synaptic efficacy of their modeled neurons and also a non-linear model introduced by (Wang and Hsu 1990) which can account for long-term habituation effects. Their neuronal model of habituation is applied on a mobile robot and habituation is observed in the robot behavior: it eventually stops responding to repeated stimuli. The model shows hierarchy of habituation (training with a stimulus of a given intensity leads to habituation to lower intensity stimuli, but not to higher intensity stimuli).

The same authors also see habituation as a mean to detect novelty by learning what is “normal” in the robot’s environment (Marsland et al. 2000). This second model is based on the Habituating Self-Organizing Map as a novelty filter using habituation, and they experiment on a Nomad 200 mobile robot exploring corridors.

(Marsland 2009) explores how these previous models, to which he adds a growing version of the Habituating Self-Organizing Map, can tackle problems like novelty detection, recency detection and temporal learning. Novelty detection can be achieved, as well as recency detection (detection of a recent prior occurrence of an event). However, temporal

¹Loria, UMR 7503, Nancy, France

²Université de Lorraine, Nancy, France

³CentraleSupélec, Metz, France

⁴INRIA Nancy Grand-Est, Nancy, France

Corresponding author:

Alain Dutech. INRIA NGE/Loria, 615 rue du jardin botanique, 54600 Villers-les-Nancy, France.

Email: alain.dutech@inria.fr

learning (“learning when the temporal context of inputs is important”) is discussed but not experimentally addressed in that work.

(Bi et al. 2017) design habituation neurons. The activity of these neurons is inhibited when they frequently receive an input vector which is close to their average input vector. In the context of deep convolution neural networks, the authors insert habituation layers between the inner product layers of the head of the network. They report an increase in performance of such “habituation” networks. (Hong et al. 2020) uses the same approach on a neuromorphic processor and shows that habituation neurons lead to more robust hardware with better aging properties.

In this paper, we use some ideas developed in these previous models (i.e. sensor memory and novelty effect) to propose new computational mechanisms inspired by habituation as described by biologists. We plug our propositions to an existing conceptual model of habits formation called Iterant Deformable Sensorimotor Medium (IDSM) (Egbert and Barandiaran 2014). We are interested in artificial agents that can acquire new behaviors and the IDSM framework specifically addresses the problem of using a simple kind of basic behaviors: habits. In the IDSM model, habits are seen as self-reinforcing repetitive patterns of behavior and its authors draw analogy from the self-organization of living processes to propose a model for habit embodiment as self-maintaining precarious sensorimotor structures that emerge from cyclic patterns (Barandiaran and Di Paolo 2014; Egbert 2018). IDSM performs well for *memorizing* new behaviors and reinforcing them so as to *replaying* them later. But it has some limitations when it comes to *generating* new behaviors. Our motivation is to add learning or adaptation mechanisms to the IDSM framework so as to open the possibility for the artificial agent to produce new behaviors.

In this context, we propose two family of extensions to the IDSM framework inspired by habituation. One is based on the idea of having a *sensor memory* which impacts the response triggered by a stimulus, especially if it is a stimulus that is continuously activated for a long time. The second approach we explore is based on the idea that *surprising events* trigger habituation by lowering the amplitude of response that would be normally triggered, and addresses brief and repeated stimuli. This paper is structured as follows. We start by recalling the biological definition of habituation and its behavioral correlates in section 2. The Iterant Deformable Sensorimotor Medium (IDSM) framework is then described in section 3. The two proposed family of models are detailed in section 4 and various experiments are presented and analyzed in section 5. We conclude the paper with a discussion on our models. The code used for the experiments of this paper is available online*.

2 Habituation

2.1 The characteristics of habituation

Habituation (see (Rankin et al. 2009) as revised description of (Thompson and Spencer 1966)) is a non-associative learning mechanism observed within numerous and various species.

The short definition of habituation is from the behavioral perspective: a decline of the organism’s response to a repeated stimulus. In this paper, we will also focus on another important characteristic of habituation, namely *spontaneous recovery*, that is, the ability of the response to recover at least partially in the absence of the stimulus. Although apparently simple, habituation involves several other characteristics as described by (McDiarmid et al. 2019) and given in table 1. Even though not directly addressed in this work, these characteristics are discussed in light of our models in section 6.

Here are some examples of relatively simple organisms using habituation. The natural expansion of the *Physarum polycephalum* (also known as the *blob*) is interrupted in the presence of a repulsive substance (quinine or caffeine). After prolonged exposure to this substance, the blob manages to diminish this repulsion, which allows it to pass through it (and potentially to access food or other sources of interest): there is habituation (Vogel and Dussutour 2016). Another example is the worm *Caenorhabditis elegans* whose escape response to vibratory stimulation diminishes with prolonged exposure (Rankin et al. 1990).

Habituation, as a non-associative learning mechanism, involves only one stimulus and the agent’s response to this stimulus. This is in contrast with associative learning, such as classical conditioning, which leads to associate a response (a conditioned response) to an initially neutral stimulus (hence becoming a conditioned stimulus).

In this study, we propose formal models that focus on the two main characteristics of habituation, namely the decrease of the response to a repeated stimulus and the spontaneous recovery.

2.2 Expected behaviors

In this part, we illustrate the behavior we expect from our models inspired by habituation using a thought experiment, illustrated on figure 1. As a reminder, we target a decrease in response when a stimulus is active and partial recovery of the response when the stimulus is absent. We use a setup with a simple agent, with only one motor and one sensor. There is one stimulus in the environment that can be either on or off. When the stimulus is active, the agent perceives it with a constant value whatever the agent’s position in the environment. The agent can move along a unique dimension, up or down. Its reflex behavior is to move up in the absence of stimulus and to go down when the agent perceives the stimulus.

In such a context, successful habituation should result in inhibiting the downward motion response after a continuous or repeated presentation of the stimulus.

In figure 1, we consider two setups and, for each, we show how the agent is expected to behave when there is no stimulus (always the top trajectory), when there is a stimulus and no habituation (always the bottom trajectory) and when there is a stimulus and habituation. When a stimulus is presented, it can be either continuous or repeated and, to

*https://gitlab.inria.fr/openbiscuit/papercode/2021_habituationgillard

1. *Habituation*: repeated application of a stimulus results in a progressive decrease in one or more parameters of a response to an asymptotic level.
2. *Spontaneous recovery*: if stimulation is stopped the response recovers at least partially.
3. *Potentiation of habituation*: repeated blocks of habituation training and spontaneous recovery lead to progressively deeper and more rapid habituation.
4. *Interstimulus interval effect*: with the same stimulus properties, high-frequency stimulation leads to more rapid and/or more pronounced response decrement and more rapid spontaneous recovery.
5. *Intensity of stimulus effect*: the less intense the stimulus, the more rapid and/or more pronounced the behavioral response decrement. Very intense stimuli may not show response decrement.
6. *Below-zero habituation*: even after the response has decremented to no response or to asymptotic level, the effects of stimulation may continue to accumulate (for example, by delaying the onset of spontaneous recovery).
7. *Stimulus specificity*: the response decrement shows some stimulus specificity.
8. *Dishabituation*: if a novel or a noxious stimulus is inserted into a series of habituating stimuli there is an increase in the decremented response to the original stimulus.
9. *Habituation of dishabituation*: if a response is repeatedly habituated and dishabituated, the magnitude of the dishabituation decreases.
10. *Long-term habituation*: some stimulus repetition protocols may result in properties of the response decrement lasting hours, days or weeks (e.g. faster rehabituation than baseline, smaller initial responses than baseline, smaller mean responses than baseline, less frequent responses than baseline).

Table 1. Characteristics of habituation from (McDiarmid et al. 2019; Rankin et al. 2009)

Illustration of habituation and spontaneous recovery

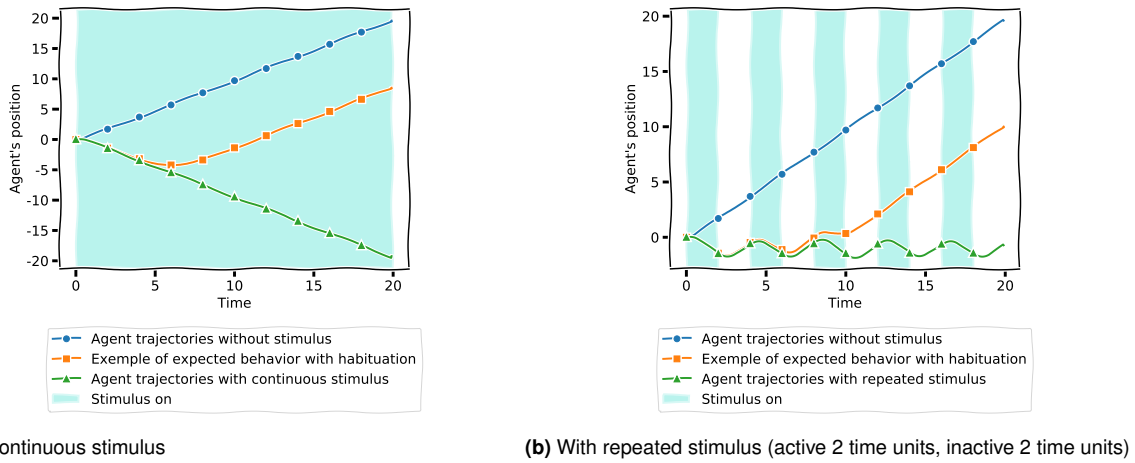


Figure 1. Illustration of the expected behaviors with or without habituation learning

ease readability, the periods when the stimulus is active are represented with a colored background.

In the first setup (Figure 1a), the stimulus is continuously on. In the second setup (Figure 1b), the stimulus is repeatedly activated for half of the time. The top (no stimulus) and bottom (no habituation) trajectories serve as a reference. An agent with habituation learning should start by following the bottom trajectory, then should gradually break away and tend toward the top trajectory, to finally be parallel to it once it reaches maximum velocity if the stimulus is totally ignored. Note that these plots are not the results of simulation but are “hand-drawn” to illustrate the properties of habituation we focus on in this paper.

3 IDSM

3.1 Formal description

The Iterant Deformable Sensorimotor Medium (or IDSM for short) of (Egbert and Barandiaran 2014) is a conceptual model of habits formation from an artificial agent perspective as self-reinforcing sensorimotor behavioral patterns. It is an intermediate model between neuronal models and macroscopic behavior models.

The IDSM works in the sensorimotor space SM of the agent. As depicted on figure 2, the environment state S_E perceived by the agent is translated into the sensorimotor state x_{SM} which comprises both sensor and

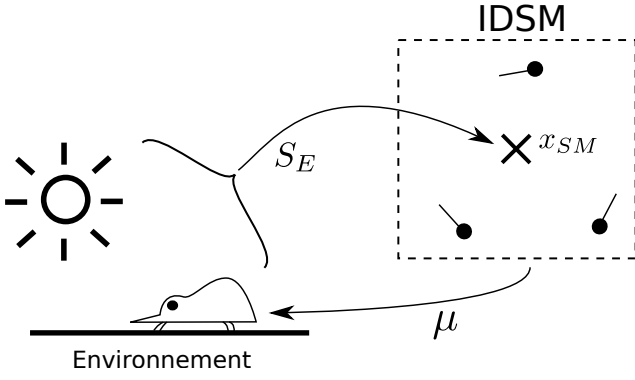


Figure 2. An IDSM works in the sensorimotor space of the agent. According to the current sensorimotor position x_{SM} and to the various nodes of the IDSM, it produces commands μ that control the agent behavior in its environment. S_E is the complete state of the environment.

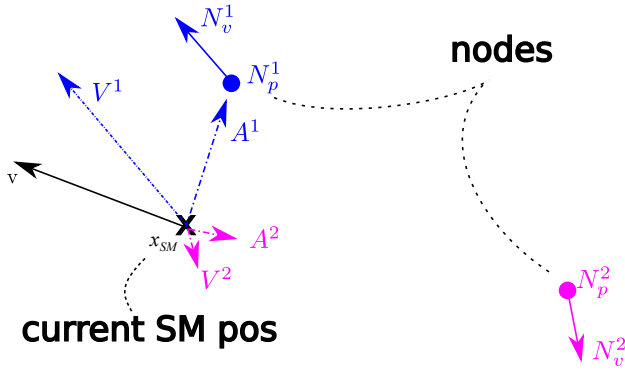


Figure 3. An IDSM is built from a collection of nodes N^i which influence the agent dynamics with a velocity V^i and attraction A^i components.

motor components. In the sensorimotor space, the IDSM uses x_{SM} and some *nodes* that store preferred sensorimotor alterations to compute a command μ . This command μ , when applied on the agent in its environment will alter the environment and the agent.

Formally, a node N of the IDSM is defined as:

$$N = \langle p, v, w \rangle \quad (1)$$

with

- p a vector indicating the node's position in sensorimotor space SM,
- v a vector indicating the desired velocity in sensorimotor space SM at the position p ,
- w a scalar indicating the node's weight, which balances its influence relatively to the other nodes within the computation of the dynamic of the agent.

The sensorimotor space is normalized and each dimension (either sensor or motor) lies in the range $[0, 1]$.

As illustrated on figure 3 the influence of each node consists of two components: an attraction component A^i and a velocity component V^i . These components impact the agent's velocity, the attraction component seeking to attract the agent towards its position N_p and the velocity component seeking to align the agent's velocity with N_v . The amplitude of the components is modulated by the distance of the agent

to the node (the closer the higher) and by the weight of the node N_w (equations (2)-(7)). The velocity command μ of the agent, normalized in the sensorimotor space SM, is given by:

$$\dot{\mu}(t) = \frac{d\mu}{dt} = \begin{cases} \frac{V(x_{SM}) + A(x_{SM})}{\phi_N(x_{SM})} & \text{if } \phi_N(x_{SM}) > \epsilon \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

with the velocity V and attraction A components defined by equations (3)-(8).

$$V(x_{SM}) = \sum_N \xi_N(x_{SM}) \cdot [N_v]^\mu \quad (3)$$

$$A(x_{SM}) = \sum_N \xi_N(x_{SM}) \cdot [\Gamma(N_p - x_{SM}, N_v)]^\mu \quad (4)$$

$$\xi_N(x_{SM}) = \omega(N_w) \cdot d(N_p, x_{SM}) \quad (5)$$

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

$$d(N_p, x_{SM}) = \frac{2}{1 + \exp(k_d \|N_p - x_{SM}\|^2)} \quad (7)$$

$$\Gamma(a, N_v) = \begin{cases} a - \left\langle a, \frac{N_v}{\|N_v\|} \right\rangle \frac{N_v}{\|N_v\|} & \text{if } N_v \neq 0 \\ a & \text{otherwise} \end{cases} \quad (8)$$

with $k_\omega = 0.025$ and $k_d = 30$. Here, $[x]^\mu$ denotes the motor projection of the a vector x in sensorimotor space. The function $\Gamma(x, y)$ keeps only the component of x which is perpendicular to y .

In equation (2), the local density of the nodes $\phi_N(x_{SM})$ is computed using:

$$\phi_N(x_{SM}) = \sum_N \xi_N(x_{SM}) \quad (9)$$

The differential equation (2) is integrated with forward Euler using $\Delta t = 0.1$.

It should be noted that, unlike the original article, we do not vary the weights over time in our experiments and we do not create new nodes to simplify the analysis of our models. Therefore we do not describe these mechanisms here, but they are available in our code as options.

3.2 Model instantiation

For our experiments, we build an IDSM with two nodes as shown on Figure 4. The two nodes are defined by:

- $N_1 = \langle p = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, v = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, w = 0 \rangle$,
- $N_2 = \langle p = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, w = 0 \rangle$.

These two nodes lead to an agent's dynamics shown by the steam plot and produces the expected behavior: if the agent perceives the stimulus (the sensory response is $s = 1$), its velocity decreases as N_1 is the most influential node; if the agent does not perceive the stimulus (the sensory response is $s = 0$), its velocity increases as N_2 is the most influential node.

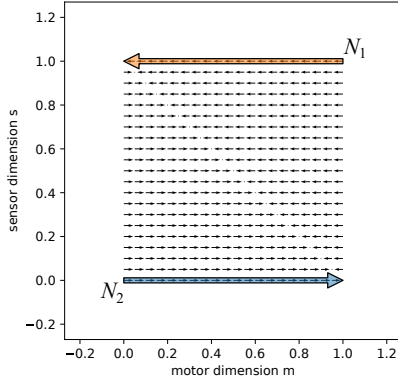


Figure 4. For our experiment, we consider an IDSM with two nodes. See text for details.

4 Models

In order to integrate habituation as defined in section 2 into the IDSM, we propose three models. In the first one (see section 4.1), a sensor memory inhibits the perception of the stimulus. In the other two, habituation occurs through the weakening of nodes influence, either by adding an energy component to the IDSM (see section 4.2.1) or through weight modification (see section 4.2.2). The model with sensor memory was designed with continuous stimuli in mind whereas the models with nodes penalty were designed to handle repeated stimuli.

4.1 Habituation through sensor memory

As habituation is defined as a decrease in response to a repeated stimulus, the first model we propose is based on a sensor memory δ , integrating the value of the sensor, and decreasing the agent's perception of the stimulus. This decrease in perception ultimately causes a decrease in the expression of the response. The sensor memory δ keeps track of the evolution of the perceived stimulus. In this model, the sensor value seen by the IDSM is the *habituated* sensor value hs , which is modulated by a function (habituation factor h) of the difference between the sensor value and its mean over a recent time window (which is the sensor memory δ). The closer the sensor memory δ is to the sensor value, the more the motor response is inhibited, without altering the nodes in the IDSM (habituation). If the stimulus is switched off, the sensor memory δ will decrease and then the motor response recovers progressively (spontaneous recovery).

The sensor memory δ is computed as a low-pass filter over the *non-habituated* sensor value nhs :

$$\frac{d\delta}{dt}(\mathbf{x}_E, t) = \beta(nhs(\mathbf{x}_E) - \delta(\mathbf{x}_E, t)) \quad (10)$$

with $\beta = 0.25$. The higher this parameter is (between 0 and 1), the higher the model sensitivity to stimulus variations and the faster it reaches asymptotic level.

The habituated sensor value hs , used as the sensor component of the agent sensorimotor position $\mathbf{x}_{SM}$, is computed from the sensor value nhs and the sensor memory

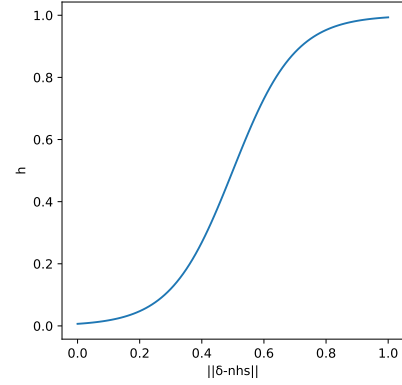


Figure 5. The habituation factor h as a function of the difference between the sensor value $nhs(\mathbf{x}_E)$ and the sensor memory δ .

δ as:

$$hs(\mathbf{x}_E, t) = h(\mathbf{x}_E, t)nhs(\mathbf{x}_E) \quad (11)$$

$$h(\mathbf{x}_E, t) = \frac{1}{1 + e^a} \quad (12)$$

$$a = \alpha \left(\|\delta(\mathbf{x}_E, t) - nhs(\mathbf{x}_E)\|_2 - \frac{1}{2} \right) \quad (13)$$

with $\alpha = -10$. From the plot of the habituation factor h on Figure 5, the stimulus is completely perceived if its value is far from the memory (large $\|\delta - nhs\|$) and completely inhibited if it is equal to the sensor memory ($\delta \approx nhs$).

4.2 Habituation through nodes penalty

While the previous model globally affects the sensory perception, it is also possible to implement habituation by local mechanisms, at the node level in the IDSM. Each node can be locally affected to produce habituation by locally inhibiting its influence on the motor response.

The two following models are based on the concept of surprise in the sensorimotor space. The idea is that each time the agent is surprised by its actual position in the sensorimotor space, the surrounding nodes are penalized. The more often a node is penalized, the less influence it gets on the IDSM dynamics. Our intuition is that the pre-existing nodes encode the response behavior and when surprised, the agent will lower the influence of this response behavior, thus leading to habituation. To allow spontaneous recovery, this influence is slightly recovered in time.

The surprise sur is an error between the prediction on the expected sensorimotor state $\hat{\mathbf{x}}_{SM}$ and the observed sensorimotor state $\mathbf{x}_{SM}$.

$$sur(t) = \|\mathbf{x}_{SM}(t) - \hat{\mathbf{x}}_{SM}(t)\|_2 \quad (14)$$

$$\hat{\mathbf{x}}_{SM}(t) = \mathbf{x}_{SM}(t - \Delta t) + \begin{bmatrix} \dot{\mathbf{x}}(t - \Delta t) \\ 0 \end{bmatrix} \Delta t \quad (15)$$

With this definition of prediction $\hat{\mathbf{x}}_{SM}$, we may observe that $[\mathbf{x}_{SM}(t)]^\mu = [\hat{\mathbf{x}}_{SM}(t)]^\mu$ and the surprise sur always equals $[\mathbf{x}_{SM}(t)]^\sigma - [\hat{\mathbf{x}}_{SM}(t)]^\sigma$, with $[\mathbf{x}_{SM}(t)]^\sigma$ the projection in the sensor space.

4.2.1 Impact on nodes energy We derive a first local mechanism by adding an energy component $N_e \in [0, 1]$ to every node N . The influence of a node on the IDSM dynamics depends on its energy level. The lower the energy, the smaller the influence. The surprise sur will decrease the nodes energy which is recovered progressively.

In practice, the velocity and attraction functions are modified in order to take into account the nodes energy:

$$V(\mathbf{x}_{SM}) = \sum_A \xi_N(\mathbf{x}_{SM}) \cdot N_e \cdot [N_v]^\mu \quad (16)$$

$$A(\mathbf{x}_{SM}) = \sum_A \xi_N(\mathbf{x}_{SM}) \cdot N_e \cdot [\Gamma(N_p - \mathbf{x}_{SM}, N_v)]^\mu \quad (17)$$

The energy of every node N is consumed depending on the surprise sur , and the proximity between the node N and the agent sensorimotor position $\mathbf{x}_{SM}$.

$$c_e(N, t) = ((1 + sur(t))^{\beta_e} - 1) \cdot \xi_N(\mathbf{x}_{SM}) \quad (18)$$

$$N_e(t + \Delta t) = N_e(t) + \gamma_e \Delta t - c_e(N, t) \quad (19)$$

where $\beta_e = 4$ adjusts the energy consumption and $\gamma_e = 0.05$ adjusts the energy recovery, and we always keep $N_e \in [0, 1]$.

4.2.2 Impact on nodes weight We derive a second local mechanism by modulating the weights of the nodes. The weight value decreases when there is surprise sur and recovers gradually over time. This model is more in line with the original IDSM.

$$c_w(N, t) = \beta_w \cdot sur(t) \cdot \xi_N(\mathbf{x}_{SM}) \quad (20)$$

$$N_w(t + \Delta t) = N_w(t) + \gamma_w \Delta t - c_w(N, t) \quad (21)$$

with weight penalty value $\beta_w = 500$ and weight recovery value $\gamma_w = 75$. The nodes weights are kept in $]-\infty, 0]$.

5 Experiments and results

In the present section, we explain the setup used to evaluate habituation and spontaneous recovery in the models presented in the previous sections, and we highlight some characteristics of these models. Then we go a little further exploring the interstimulus interval effect and the results of the combination of the models.

5.1 Habituation and spontaneous recovery experimental setup

In order to measure if our models are capable of habituation and spontaneous recovery, we use the following protocol.

The experiment is split into different phases: stimulation phases (with continuous or repeated stimulus) and quiet phases (without any stimulus). First there are two stimulation phases to evaluate habituation, and then a quiet phase followed by a stimulation phase to evaluate spontaneous recovery. Each phase lasts twenty time unit. To limit the differences in experimental conditions, at the beginning of each phase, the agent is reset to a null velocity starting position ($\mathbf{x}_E = \dot{\mathbf{x}}_E = \ddot{\mathbf{x}}_E = 0$) and a neutral sensorimotor

state $\mathbf{x}_{SM} = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix}$. Here, the choice of forcing the sensor value to 0 is made to have identical initial conditions between phases, in particular to compare the first and fourth phases. This can be interpreted as the insertion of a time unit without stimulus between each phase.

In a given experiment, the stimulation phase follows one of the four patterns below. Note that for one experiment, all the stimulation phases follow the same pattern.

- (a): continuous stimulus;
- (b): repetitions of 3 time units on / 1 time unit off;
- (c): repetitions of 2 time units on / 2 time units off;
- (d): repetitions of 1 time unit on / 3 time units off.

5.2 Habituation and spontaneous recovery results

Habituation and spontaneous recovery are evaluated by comparing the behaviors of a learning agent with a baseline agent without learning mechanisms. The behaviors of the agents, for each model, are given on figures 6-8.

Each figure is composed of four subplots, one for each stimulation pattern. Each subplot has the same layout. The top plot displays the learning agent trajectory and the control agent trajectory. On the trajectory of the learning agent, there are two numbers: the agent last position $\mathbf{x}_E$ for this phase and a score relative to the control agent. Two scores are computed using the control agent last position $\mathbf{x}_c$ and its “best” last position $\mathbf{x}_c^*$ (the highest position reached by the control agent in the absence of stimulation). z_S is the score for the stimulation phase, it measures the efficiency of habituation and spontaneous recovery and is defined as:

$$z_S(\mathbf{x}_E) = \frac{\mathbf{x}_E - \mathbf{x}_c}{\mathbf{x}_c^* - \mathbf{x}_c} \quad (22)$$

This score lies in the range $[-\infty; 1]$, with $z_S = 0$ if the agent behaves like the control agent and $z_S = 1$ if the agent is not influenced by the stimulus, hence exhibits habituation. If the response to the stimulus is amplified, the score z_S is negative which we call *reverse habituation* (i.e. habituation to the deactivation of the stimulus, hence to the upward behavior in place of the downward behavior).

z_Q is the score for the quiet phase, it measures the lasting effect of *reverse habituation* and is defined as:

$$z_Q(\mathbf{x}_E) = -(1 - \frac{\mathbf{x}_E}{\mathbf{x}_c^*}) \quad (23)$$

This scores lies in the range $[-1; 0]$, with $z_Q = 0$ if the agent instantly behaves as the control agent and a negative z_Q reflects a lasting reverse habituation effect.

The second plot from the top gives information regarding perception of the stimulus by the agent displaying the stimulation value and the normalized sensor value of the agent in the sensorimotor space (the second component of $\mathbf{x}_{SM}$). The third plot from the top shows the normalized motor value of the agent in the sensorimotor space (the first component of $\mathbf{x}_{SM}$), and the velocity command μ . The last plots give more information depending on the model (sensor memory, energy of the nodes, weights of the nodes).

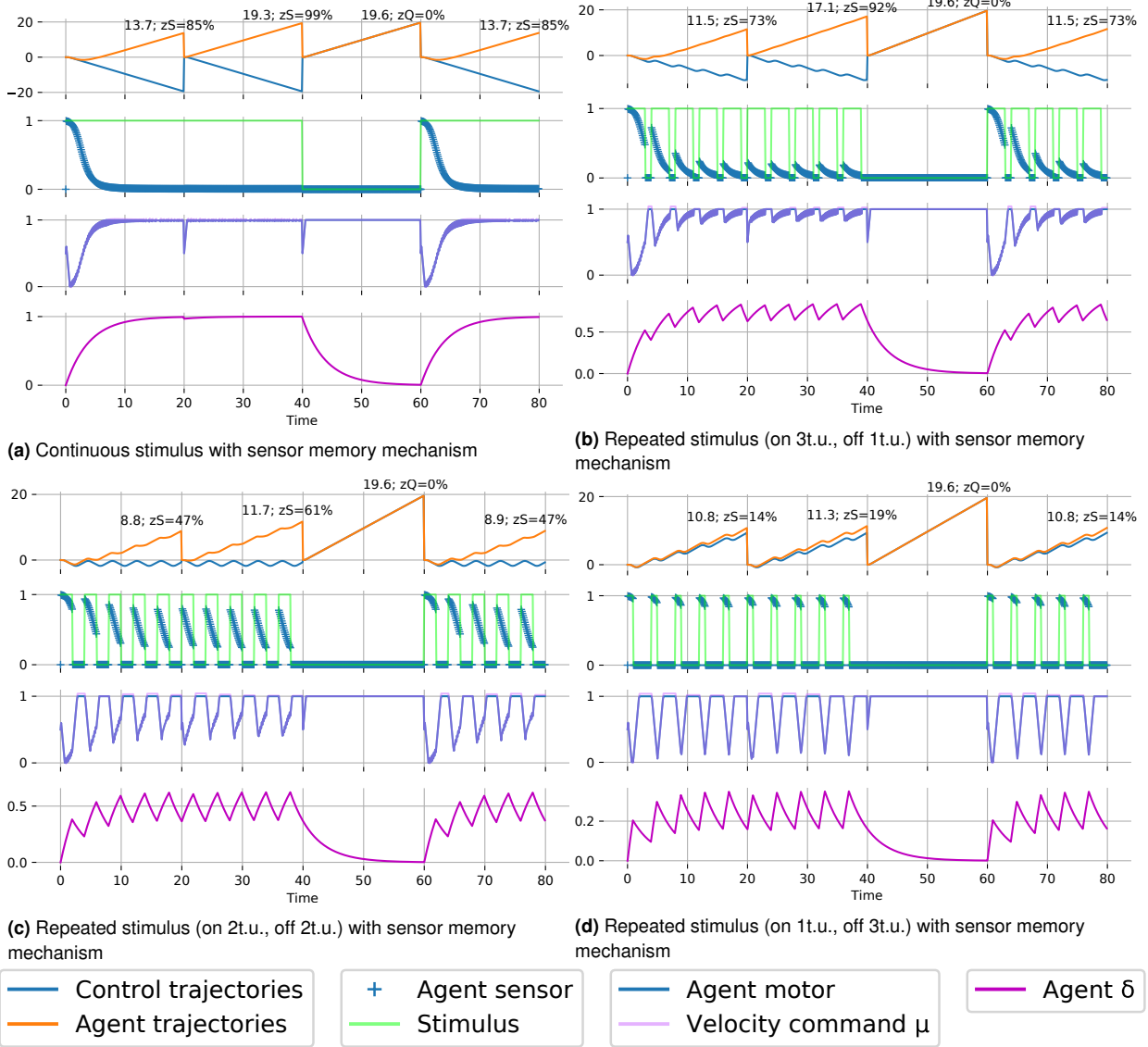


Figure 6. Sensor memory mechanism for four different stimulation pattern. For each subplot, from top to bottom: the trajectory of the learning and control agents, the stimulus and agent's sensor values, the normalized motor value and velocity command μ , the agent sensor memory δ . In each of the four cases, the learning agent quickly habituates and stops expressing the motor response (going downward) to the stimulus and moves upward. The stimulation pattern impacts the speed and amplitude of habituation.

For the control agent, a response behavior (downward motion) is indeed observed with the activation of the stimulus (phases 1, 2 and 4). During the quiet phase, we observe the default behavior of the control agent (moving upward).

Each of the three proposed models show habituation for at least one type of stimulation pattern.

For the sensor memory model (section 4.1), as shown on figure 6, longer stimulus activation produces stronger habituation. This can be explained by the fact that the sensor memory δ approximates the mean of the stimulus activation (*i.e.* 1, 0.75, 0.5 and 0.25 respectively with stimulation patterns (a), (b), (c) and (d), see bottom plots of the four subplots), hence its inhibition effect is accordingly stronger with longer activation time (see equation (12) and figure 5).

Conversely, nodes penalty models (sections 4.2.1 and 4.2.2) exhibit a stronger habituation effect when the

duration of activation is strictly lower than the duration of inactivation (*i.e.* with stimulation pattern (d), see figs. 7 and 8). For both models, the habituation effect depends on the occurrence of surprising events. For the stimulation pattern (a), there are too few surprising events to trigger habituation. For the repeated stimuli, the ratio between activation and inactivation duration impacts habituation because it changes the interval covered by the motor component in the sensorimotor space.

To illustrate that phenomenon, we consider the sensorimotor trajectories of an agent with habituation learning with a high (figure 9c) or low (figure 9d) ratio of stimulus activation. In the case of a high ratio of stimulus activation, the agent spends more time in the top of the *SM*, under the influence of node N_1 which brings the agent's sensorimotor motor component close to

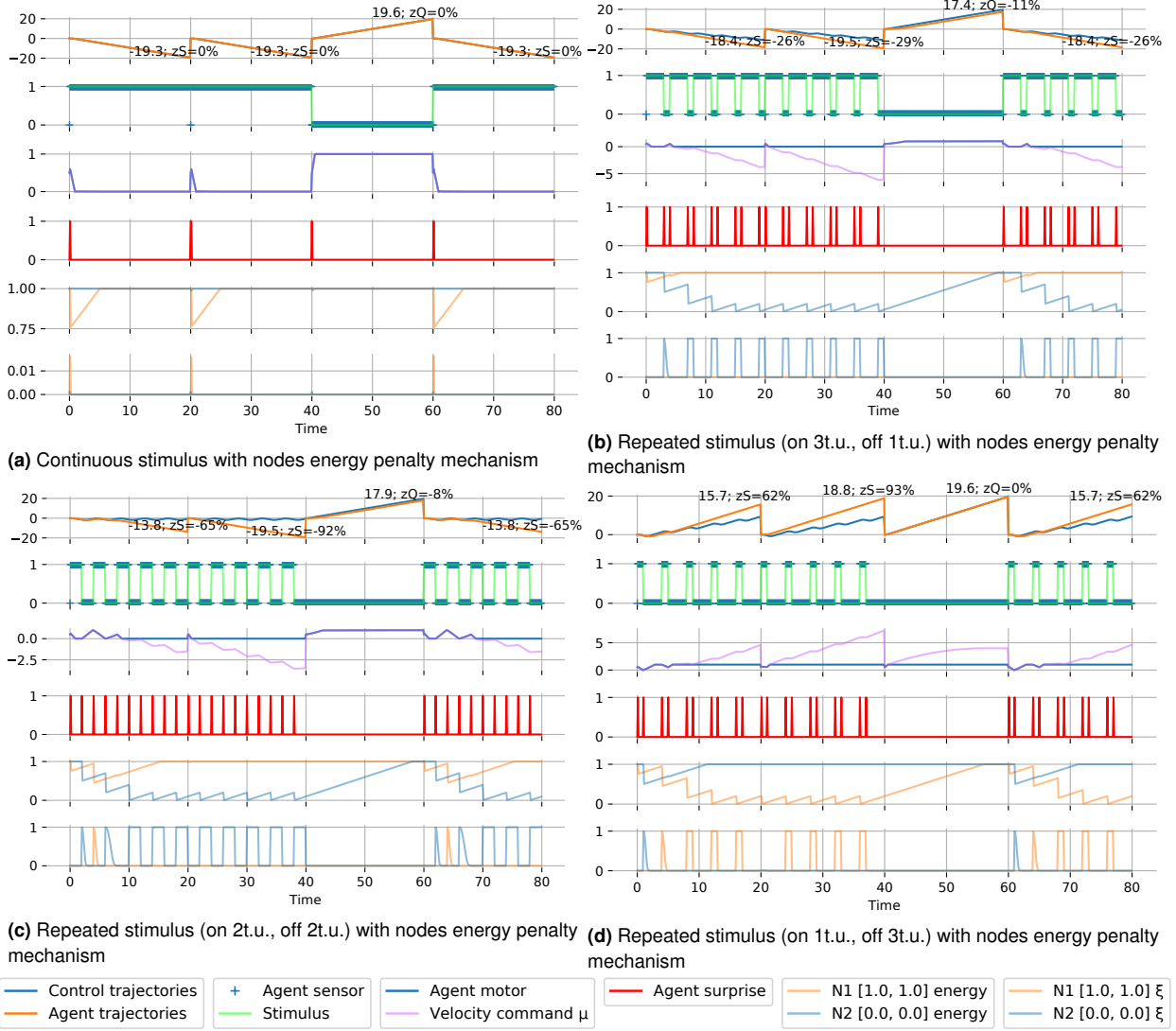


Figure 7. Nodes energy penalty mechanism for four stimulation patterns. For each subplot, from top to bottom: the trajectory of the learning and control agents, the stimulus and agent's sensor values, the normalized motor value and velocity command μ , the surprises, the nodes energy and the nodes ξ_N . In the first plot (a), the lack of surprise leads to no habituation contrary to the last plot (d) where habituation takes places. In plots (b,c), we observe reverse habituation (because the timing of surprising events leads to a decrease of N_2 energy).

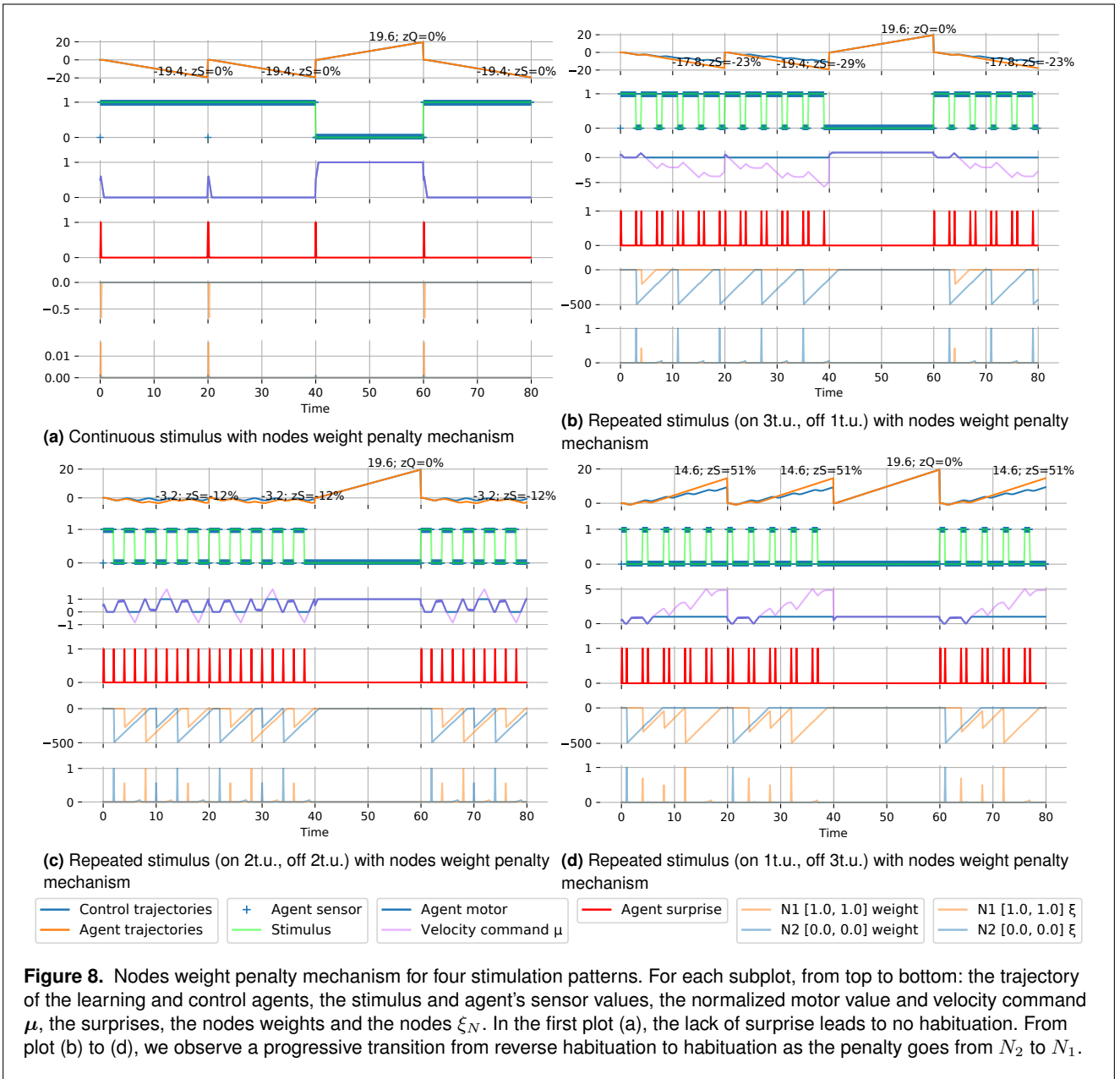
0. Therefore, when the stimulus is deactivated, the agent goes straight on the node N_2 , which takes full penalty. The weakened N_2 influence combined with the shorter inactivation time reduces the distance covered by the agent before reactivation of the stimulus. Hence, even when the stimulus is reactivated, the agent is never close enough to N_1 for it to be penalized. Progressively, N_2 is penalized, N_1 is unaffected and asymptotically, the agent oscillates in sensorimotor space between positions with a null motor component. This corresponds to a *reverse* habituation situation where the default behavior (upward) is inhibited.

In Figure 9d with low ratio of stimulus activation, the three first steps are the same. Then, although the node N_2 undergoes a full penalty on its first “visit” by the agent, the inactivation time is longer and sufficient for the agent to go to $m = 1$. The same arguments than previously with opposite conditions (shorter activation time and strong penalization of N_1) explain why the agent asymptotically oscillates

in sensorimotor space between positions with a motor component equals to 1. This corresponds to the habituation where the response behavior (downward) is inhibited.

Now let's consider the spontaneous recovery characteristic. For the three models, evaluation of the spontaneous recovery involves the last two phases of the protocol (the quiet phase and the last stimulation phase). Spontaneous recovery is expected to occur during the quiet phase and its impact is evaluated during the following stimulation phase (phase 4). Indeed, the agents trajectories during their fourth phase are similar to those in their first phase. This is explained by the fact that by the end of third phase, all models core components (sensor memory, nodes energy or nodes weight) return to values close to those in the first phase.

As a conclusion of this section, we note that all the three proposed models do exhibit habituation and spontaneous recovery. The stimulation ratio impacts the amplitude of the habituation. This impact is not the same for the three models.



In particular, continuous activation only triggers habituation in the sensor memory model. Finally, it is worth noting that the sensor memory model, through β , can adjust the speed of habituation and of spontaneous recovery, but not separately. The higher β is, the faster both habituation and spontaneous recovery are. On the nodes penalty models, the two characteristics can be regulated independently by adjusting independently their γ (to regulate habituation) and β (to regulate spontaneous recovery) parameters.

5.3 Interstimulus interval setup and results

We focus here on the interstimulus interval effect which is the fourth characteristic of habituation as described in (Rankin et al. 2009). According to biologists, there are three potential consequences of higher frequency stimulation: more rapid response decrement, more pronounced response decrement and more rapid spontaneous recovery.

We test this characteristic by using a new protocol with two phases. During the first phase, we activate the stimulus ten times, each activation followed by an interstimulus interval, and then we have a quiet phase to evaluate the spontaneous recovery. The interstimulus interval value is set either to 2, 3, 4 or 5 time units. The stimulus duration is always of 1 time unit.

With the sensor memory model (see figure 10), the agent behavior shows a more pronounced response decrement when there is a higher stimulation frequency (score $z_S = 43\%$ at the highest frequency, $z_S = 10\%$ at the lowest). With this protocol, it is not easy to tell if it is more rapid or not from a behaviour perspective. Since we have access to some internal values, our analyses can go beyond this behavioural level. Looking at the sensor memory δ , from highest to lowest frequencies, it takes respectively 15.9, 14.8, 14.1 and 13.6 time units after the last deactivation of the stimulus for δ to go from respectively 0.56, 0.42, 0.35 and 0.31 to below 0.01. Thus with higher frequencies,

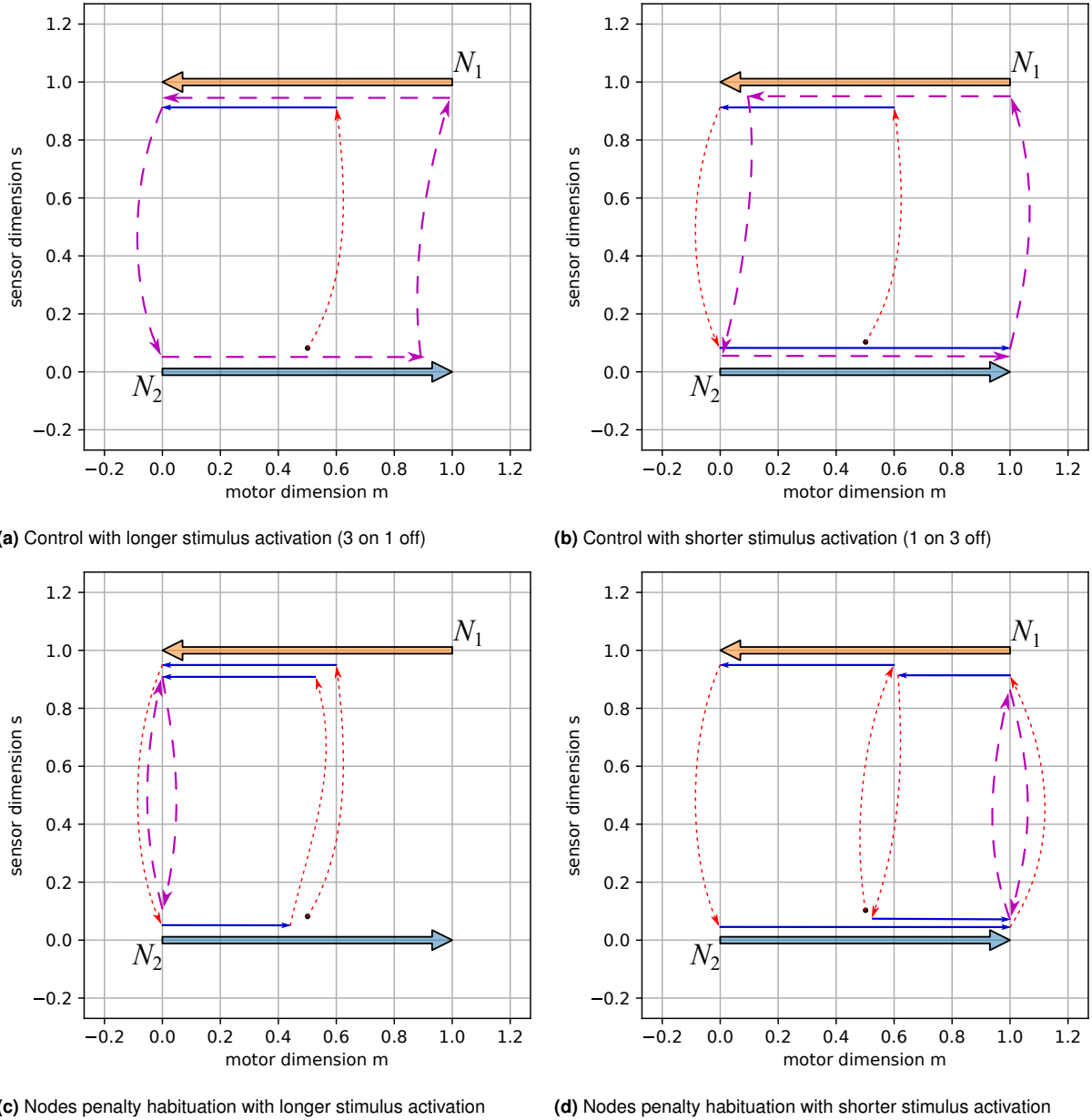


Figure 9. Sensorimotor trajectories showing the impact of the ratio of stimulus activation and inactivation. Top and bottom thick arrows represent the two nodes. Horizontal thin arrows represent the agent trajectory in the sensorimotor space, and should always either be at $s = 1$ when the stimulus is on, or at $s = 0$ when the stimulus is off, but are spaced for better readability. Vertical curved arrows show jump in the trajectory due to activation or inactivation of the stimulus. Arrows with big dashes represent the cycle in which the agent ends. The agent always starts at the black dot (also vertically spaced for readability). Top plots depict trajectories of the control agent, bottom ones are trajectories of agent with a nodes penalty model. Bottom right illustrates habituation whereas bottom left illustrates *reverse* habituation.

the spontaneous recovery takes longer, but its rate is faster: respectively 0.035, 0.028, 0.024 and 0.022. This recovery rate r is calculated according to the following formula:

$$r = \frac{\delta_1 - \delta_2}{t} \quad (24)$$

with δ_1 the value of δ after the last deactivation of the stimulus, δ_2 the first value below 0.01 and t the time needed for the decrease between δ_1 and δ_2 . The general tendency of δ is to increase until it oscillates around its asymptotic mean value. This mean value is higher for higher stimulation

frequencies. The interstimulus interval values are always too short for δ to return to 0, so it accumulates during every stimulus activation. The amount accumulated during the one time unit of activation depends on the difference between the start value and the asymptotic mean value. A crude way to evaluate the more rapid habituation effect is to measure the number of activation cycles or the time needed for δ to reach its asymptotic mean value. In both cases, higher frequencies lead to less rapid habituation with this model: respectively, from higher to lower frequencies, 8.8, 6.7, 4.7 and 0.9 time units needed to go to the respective mean values (0.5, 0.33,

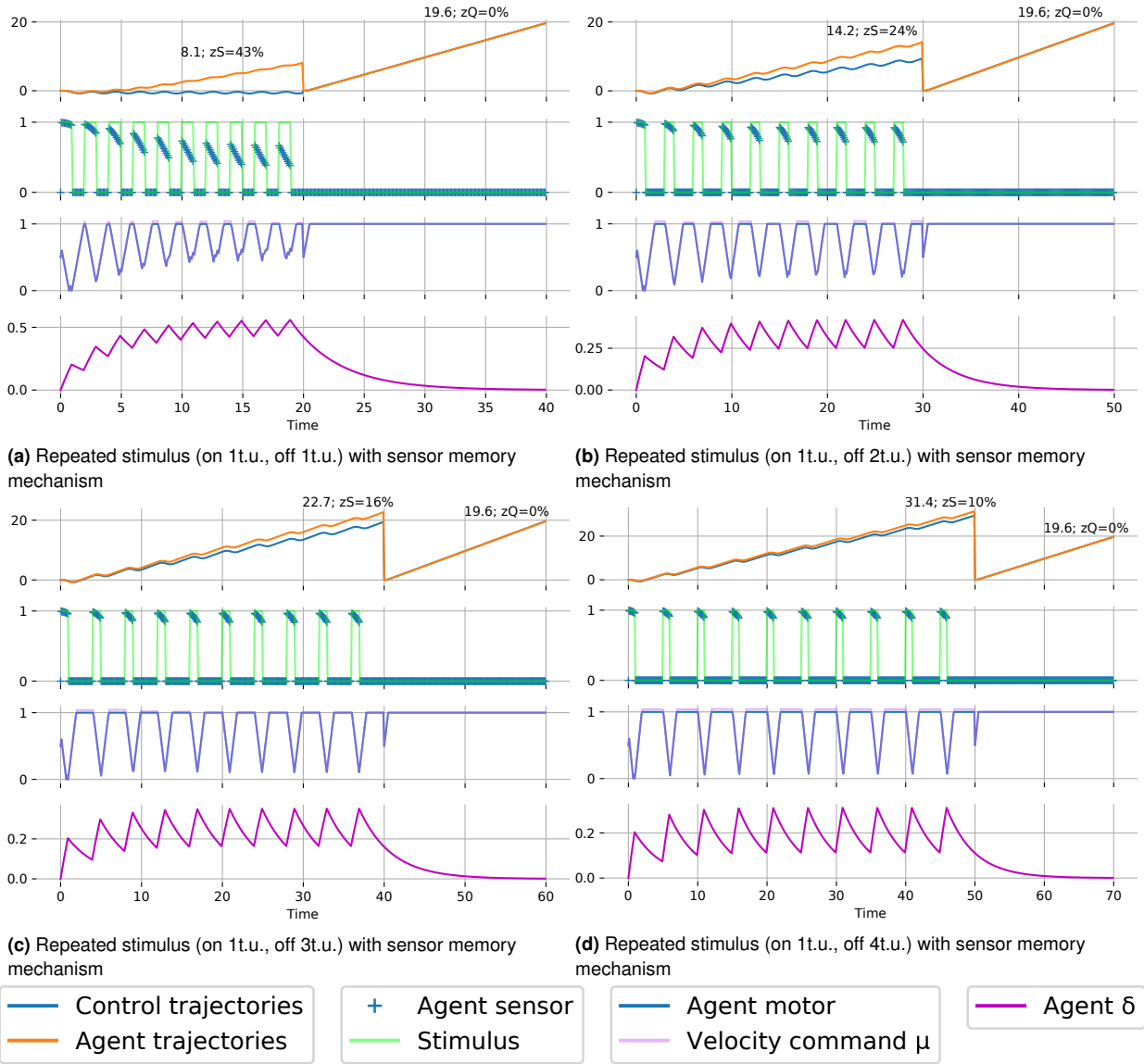


Figure 10. Test of the interstimulus interval effect with the sensor memory mechanism. For each subplot, from top to bottom: the trajectory of the learning and control agents, the stimulus and agent's sensor values, the normalized motor value and velocity command μ , the agent sensor memory δ . Habituation decreases as the interstimulus interval increases, keeping a constant activation time. Note the duration of the experiments vary from 40 t.u. to 70 t.u. .

0.25 and 0.2), hence fifth, third, second and first activation cycle.

For the nodes energy penalty model (see figure 11), except the unique case of *reverse* habituation in stimulation pattern (a), there is a more rapid response decrement, but not a significantly more pronounced response decrement since the score z_S is roughly equals to 80%. Indeed, with interstimulus intervals of figures 11b, 11c and 11d, the asymptotic level of the node N_1 energy is reached at the fourth activation of the stimulus. It is more rapid since this fourth activation occurs earlier with shorter interstimulus intervals. In the quiet phase, the spontaneous recovery is never more rapid, it is exactly the same (*i.e.* 20 time units) with all interstimulus intervals. It was expected as the energy recovery rate is constant.

In the case of the nodes weight penalty model (see figure 12), with decreasing frequency the agent goes from *reverse* habituation ($z_S = -10\%$) to more and more

pronounced habituation ($z_S = 75\%$ at the lowest frequency). In the quiet phase, the spontaneous recovery is never more rapid, as the weight recovery rate is constant. The rapidity aspect of the response is more complex to analyse. The lower the frequency, the more the sensorimotor trajectory of the agent is confined to the right side (motor component close to 1) of the sensorimotor space. The node N_1 is inhibited earlier in the case of high frequency, but the penalty is stronger in the case of low frequency. The global effect on the speed of response decrement is difficult to evaluate.

Overall, our models never show the complete interstimulus interval effect. The more rapid spontaneous recovery is never displayed, except if we consider its faster rate with the sensor memory model. In fact, in our models, the interstimulus interval effect depends mostly on the spontaneous recovery, and in particular on how, or when, it

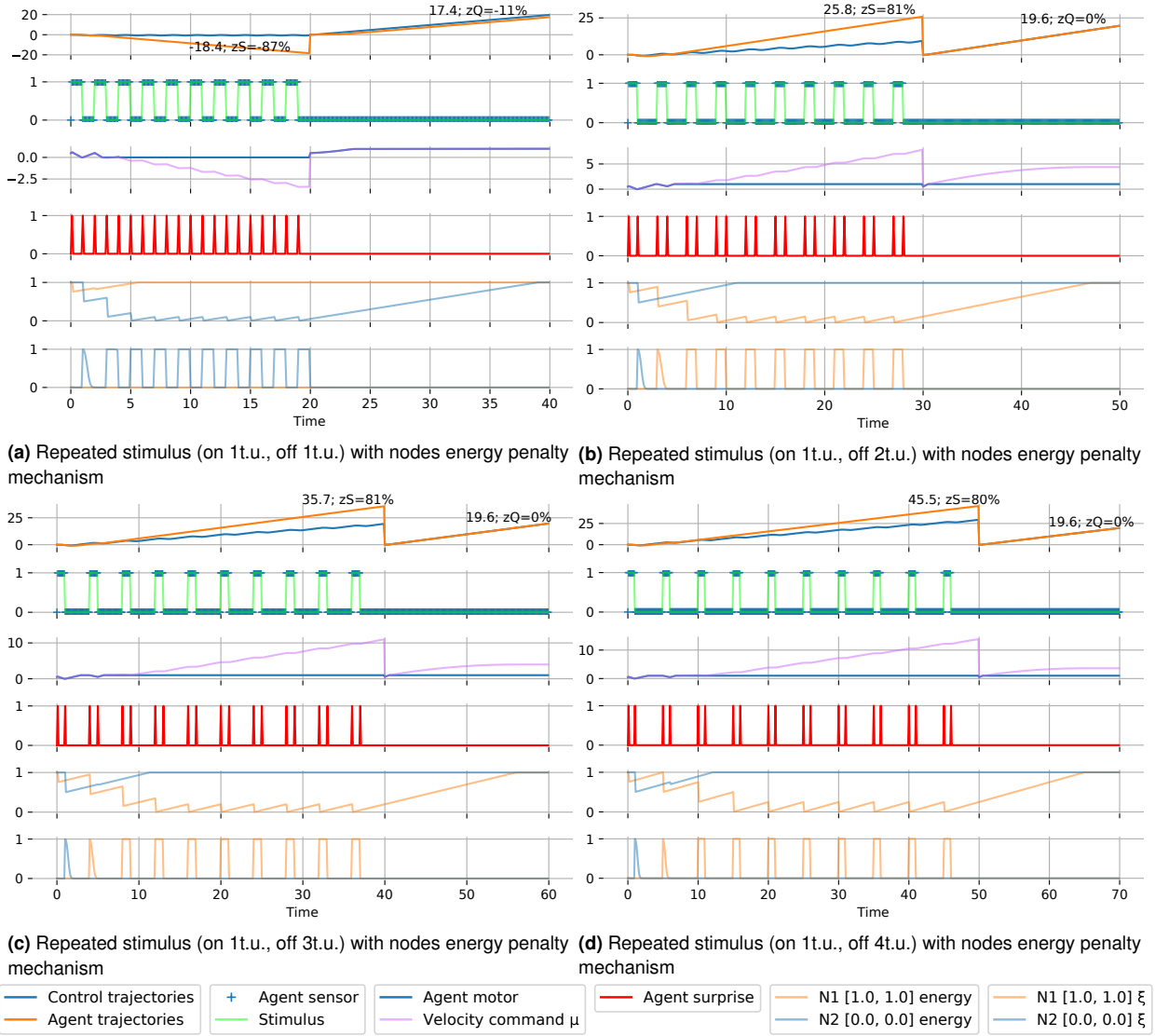


Figure 11. Test of the interstimulus interval effect with the nodes energy penalty mechanism. For each subplot, from top to bottom: the trajectory of the learning and control agents, the stimulus and agent's sensor values, the normalized motor value and velocity command μ , the surprises, the nodes energy and the nodes ξ_N . For plot (a), we observe reverse habituation. From plots (b) to (d), as interstimulus interval increases, habituation is less rapid to setup. Note the duration of the experiments vary from 40 t.u. to 70 t.u. .

takes place: permanently (nodes penalty models), or directly after the last perceived stimulus (sensor memory model).

5.4 Combination of models

Each of our models have its own properties, which allow the agent to habituate more or less effectively depending on the stimulation pattern (see section 5.2). But these simple models do not solve all cases: with fixed parameters, there will always be a situation where the habituation will be more difficult to achieve. This section explores some combination of our models to obtain habituation with all the stimulation patterns. These combinations are all the easier to achieve as the habituation mechanisms are complementary since they have an impact on components that are independent of each other.

We combine the sensor memory mechanism with each of the nodes penalty models (see figs. 13 and 14). The

first observed result is that there is no loss: the habituation resulting of such combination is always equal or better than the habituation of the same models used individually. Indeed, the habituation score z_S of figs. 13 and 14 are always greater or equal than the habituation scores of figs. 6 to 8.

Besides, in the cases where a nodes penalty mechanism alone would result in *reverse* habituation (negative z_S , with stimulation patterns (b) or (c)), the combined model does result in habituation. This is mainly explained by an asymmetry in the surprise value which is bigger at the activation of the stimulus than at its deactivation. In our definition of surprise (eq. 14) and implementation of the predicted next state (eq. 15), and as illustrated on fig. 15, surprise equals the difference between two successive sensor value (represented by cross on the figure). By building up during stimulus activation, the sensor memory δ causes a decrease in the habituated sensor value (represented by

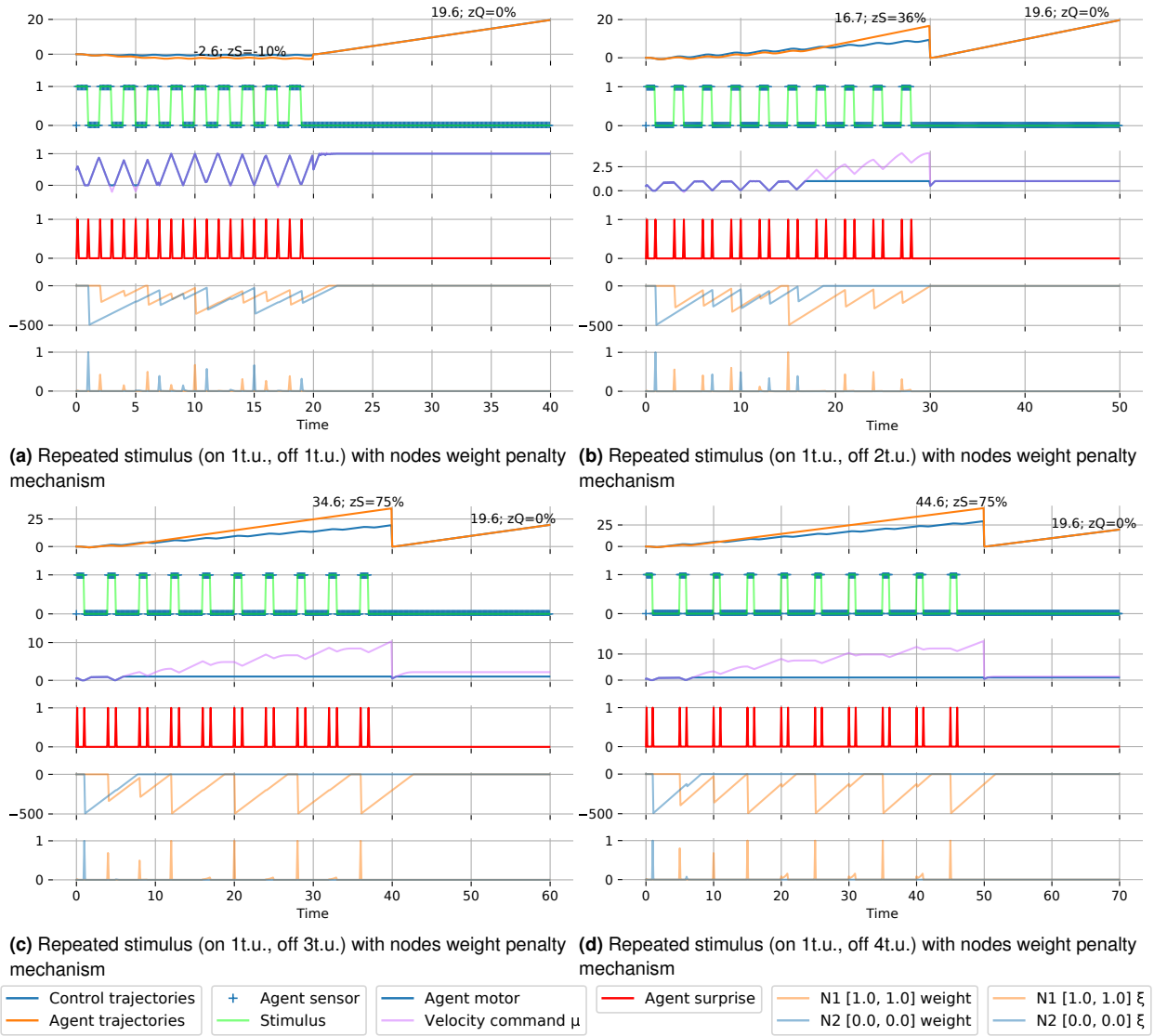


Figure 12. Test of the interstimulus interval effect with the nodes weight penalty mechanism. For each subplot, from top to bottom: the trajectory of the learning and control agents, the stimulus and agent's sensor values, the normalized motor value and velocity command μ , the surprises, the nodes weights and the nodes ξ_N . For all plots except (a), habituation takes place but relies on a more subtle interaction between the N_1 and N_2 penalties than the other experiments. See the main text for more details. Note the duration of the experiments vary from 40 t.u. to 70 t.u. .

cross). As a consequence, surprise value is lower at stimulus deactivation than at stimulus activation. Node N_2 is less inhibited than node N_1 because node N_2 is penalized by the weaker *deactivation* surprise while node N_1 is penalized by the stronger *activation* surprise. As a result, the node penalty mechanism favors habituation in place of *reverse* habituation.

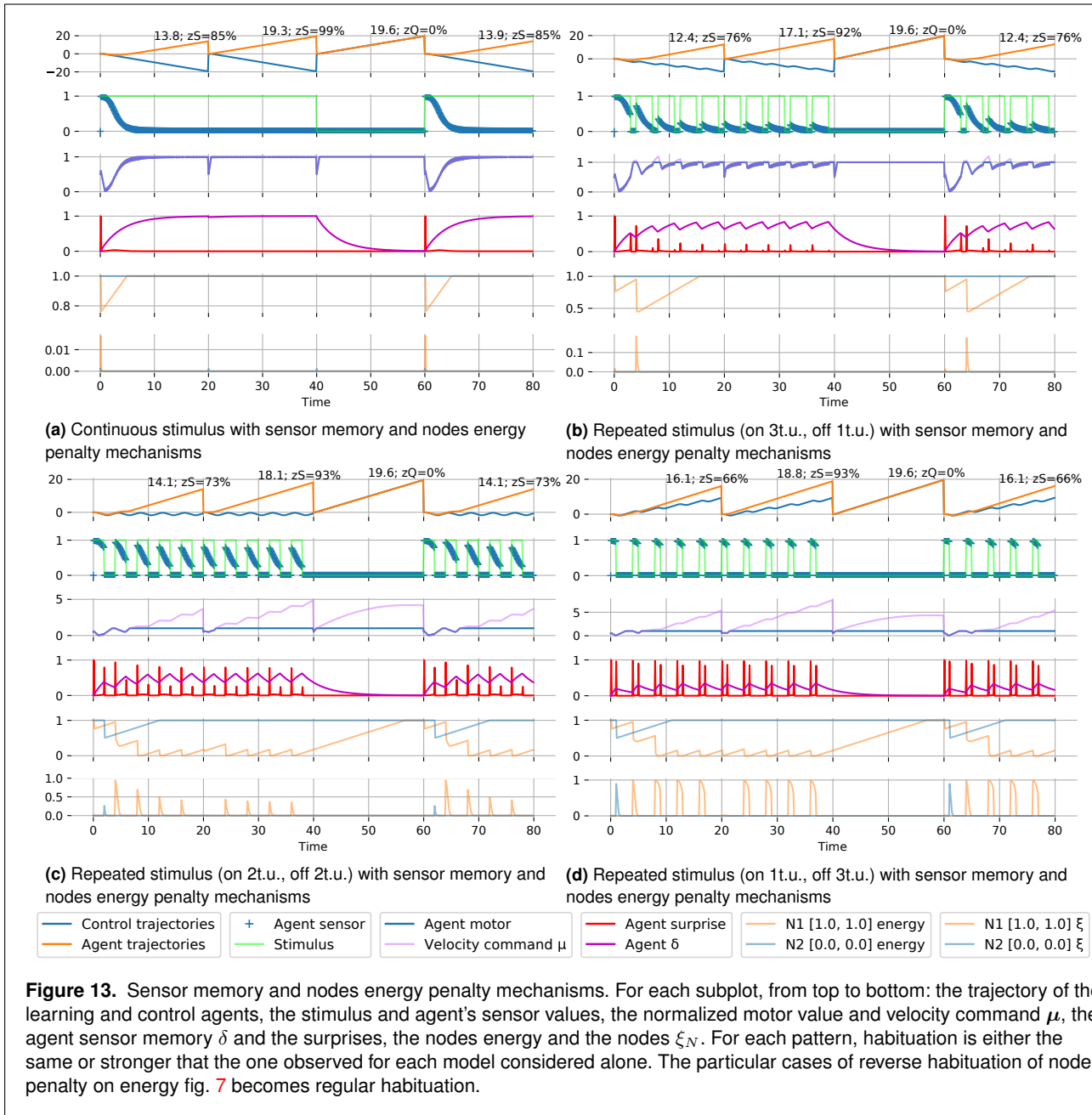
6 Discussion

In this section, we first discuss some differences between our models. Then we give our thoughts about the characteristics of habituation not covered in the result section and whose study could be interested along two lines: either for modelers who would like to extend our work or because those properties might also be fundamental for behaving agents.

6.1 Models comparison

Our models are designed at the abstract level of the IDSM. In this, they are different from most of the related work cited in the introduction (Marsland et al. (1999), Wang and Hsu (1990), Bi et al. (2017), Hong et al. (2020)) that work on the level of neurons or synapses. One slight advantage is that our models, being more conceptual, can be easier to analyze and understand, but, as discussed now, an in depth understanding of our models can still be complex. On the other hand, many of these related works are more advanced than ours in term of their practical application : contrarily to (Marsland et al. 2000) we are still far from using our IDSM with habituation in a robotic setting.

6.1.1 Sensor memory vs nodes penalty The different models we propose were designed to allow habituation in response to different stimuli: the sensor memory to



specifically address a continuous stimulus and the nodes penalty to specifically address a repeated stimulus. But some differences go beyond this.

To begin with, while the nodes penalty models impact the sensorimotor dynamics (the velocity field shown for example on fig. 4), the sensor memory model alters the sensorimotor state with no modifications on the nodes (hence no modifications on the sensorimotor dynamics). With the sensor memory, we artificially change the sensor position of the agent in the sensorimotor space, which then implies a change in the motor response. In a nodes penalty model, first the nodes influence is impacted, then this affects the dynamics, which in turn affects the motor response.

Another difference is that with the sensor memory model, habituation occurs depending on the stimulus intensity value, whereas for nodes penalty models, with the surprise mechanism, habituation occurs depending on the change in

the stimulus intensity. Therefore, the nodes penalty models can account for habituation in case of a sudden decrease of stimulus intensity, as observed for example with Zebrafish larvae which show habituation to sudden decreases in illumination (see (Randlett et al. 2019)), while the sensor memory model alone could not.

It is also that inability of the nodes penalty models to be sensitive to the absolute value of the stimulus intensity that allows them to produce what we called *reverse* habituation. Another consequence for the nodes penalty models is that spontaneous recovery can happen also when the stimulus is still active (and not only when the stimulus is inactive). In this very case, these two features can be seen as a disadvantage because they prevent habituation to occur. But, having at our disposal several different mechanisms is also an advantage. For example, this reverse habituation could

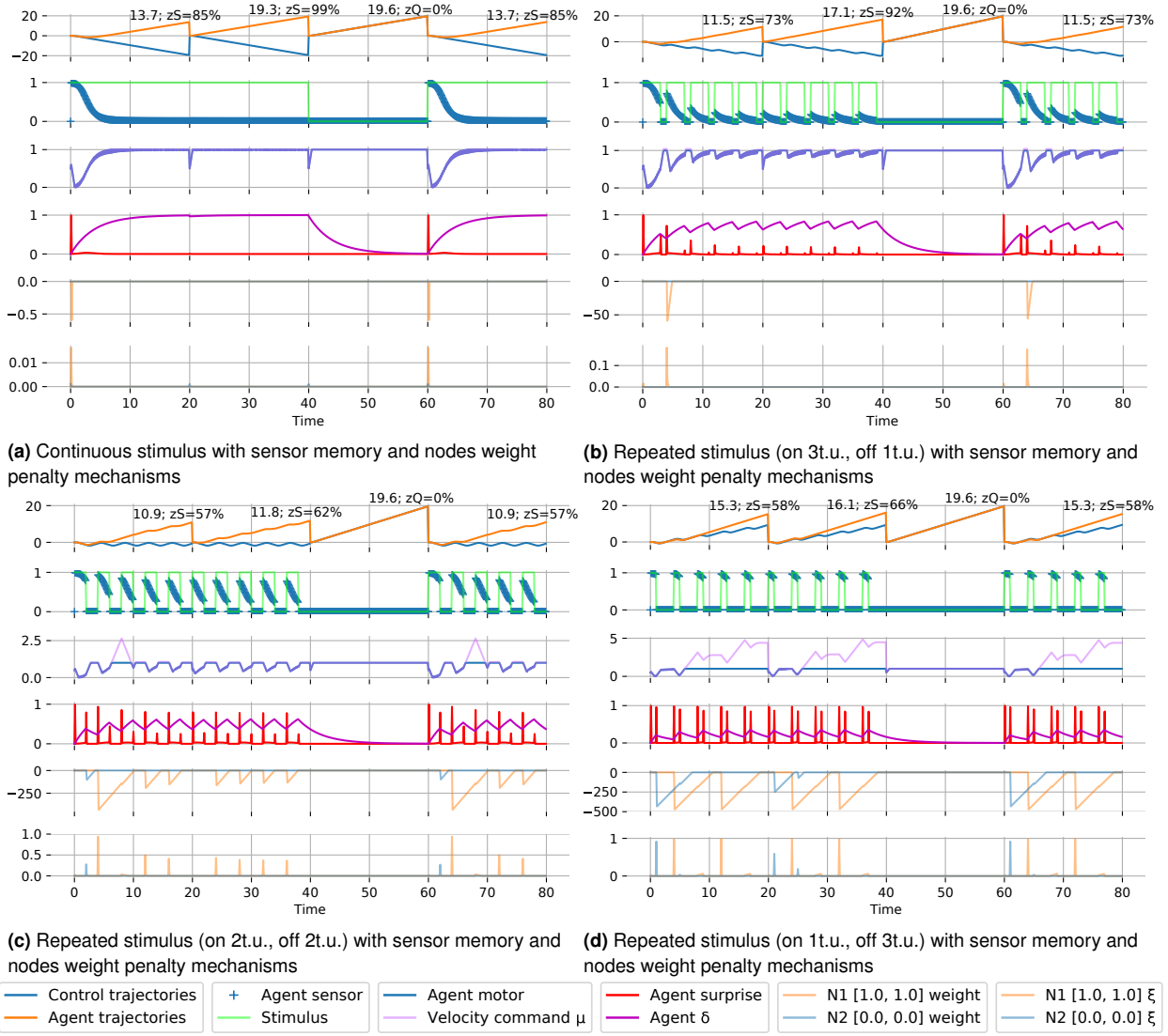


Figure 14. Sensor memory and nodes weight penalty mechanisms. For each subplot, from top to bottom: the trajectory of the learning and control agents, the stimulus and agent's sensor values, the normalized motor value and velocity command μ , the agent sensor memory δ and the surprises, the nodes weights and the nodes ξ_N . In every case, combined mechanisms lead to stronger habituation effect than any one of the single mechanism, and previous reverse habituation (of fig. 8) is no longer present.

become a kind of habituation when stimuli goes from “on” to “off”.

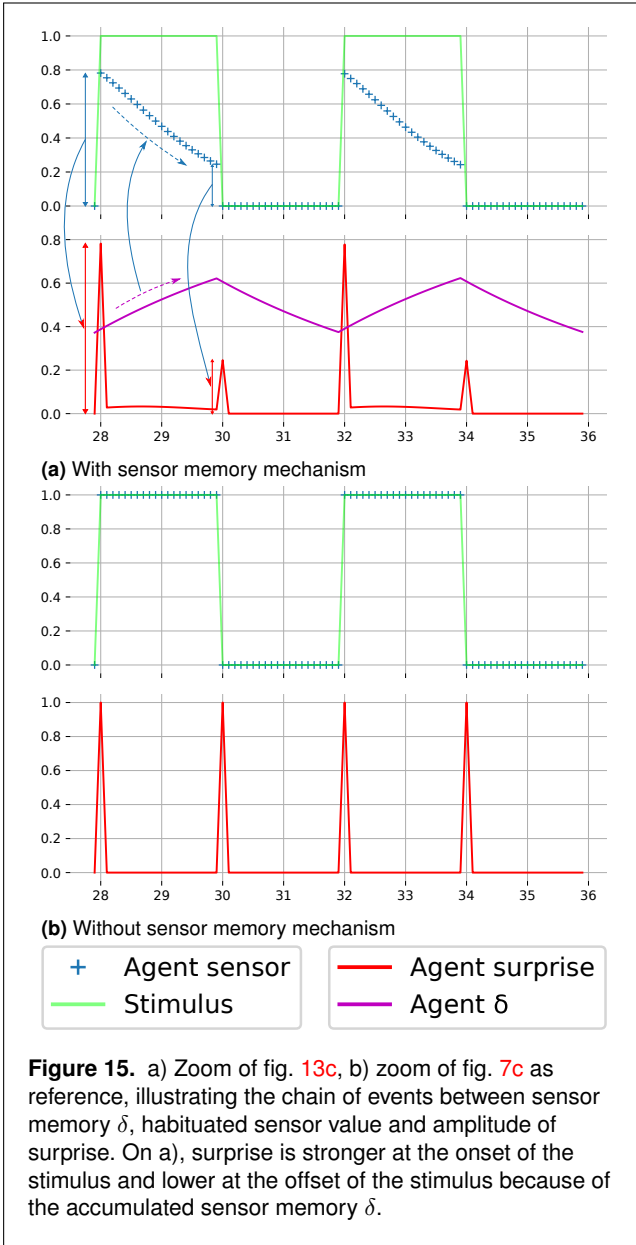
In the nodes penalty models, the motor component of the agent sensorimotor state has a very strong influence on the amplitude of the habituation effect, while it has absolutely none for the sensor memory model. It is not obvious for us if in biology habituation learning only rely on the sensory input or if it also depends on what the organism is doing. We are considering the possibility to only use the sensor components in the nodes penalty computation in future works.

Lastly, as said in the end of section 5.2, the sensor memory model can only adjust the speed of habituation and the speed of spontaneous recovery together, with a unique parameter, whereas the nodes penalty models can do so separately, through two parameters.

6.1.2 Nodes energy vs nodes weight There are also notable differences between the two nodes penalty models.

With the current parameters, energy loss is more progressive than weight loss, but nevertheless, the loss of influence of energy nodes is more important than the loss of influence of weight-penalized nodes. As explained below, it is mainly due to the distance at which a node can have an influence (a distance which is reduced when the weight is decreased), because the distance at which a node can be penalized by a surprising event is its influence distance.

The main difference is that in the model impacting nodes weight, a lower weight means a lower ω (see equation (6)). One of the consequences is that a decrease in ω leads to a decrease of the impact of the surprise at a given distance (see eqs. (20) and (21)). Therefore, the lower the node weight, the harder it is to inhibit it further, since to do so the agent sensorimotor position x_{SM} needs to be closer and closer to the node's position N_p . In the model impacting the node energy, since the energy mechanism is completely

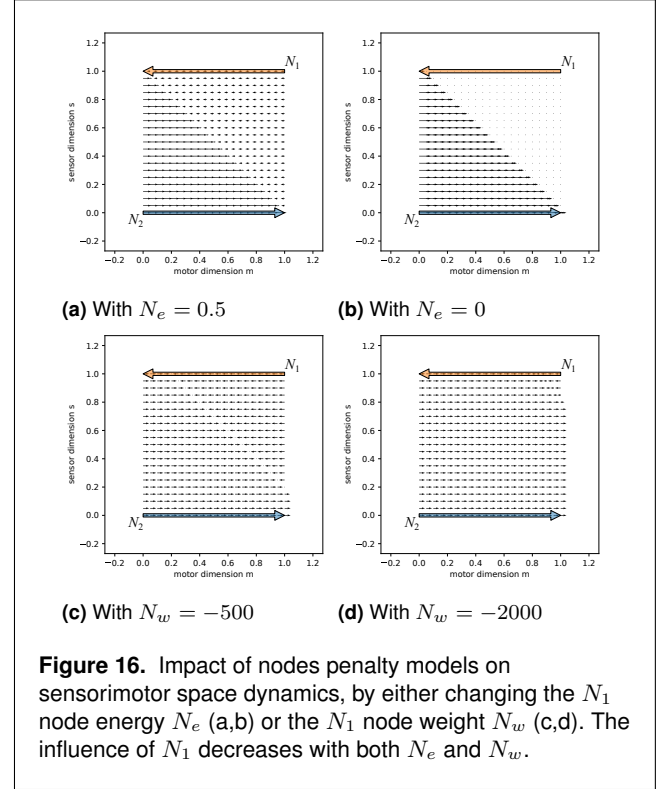


independent from the area of influence of the nodes, there is no such phenomenon: the consumption c_e and the agent sensorimotor position x_{SM} being equal otherwise, a node N will undergo the same energy penalty whether N is weakened or not. In addition to that, for a given surprise, the penalty will still be significant even at a greater distance in the model impacting energy than in the one impacting the weights, even at initialization (see for example first penalties in the fifth plot from the top in figs. 7a and 8a).

Another significant difference is that the model impacting nodes energy decreases the node influence whereas the model impacting nodes weight decreases the node influence *range*. With decreasing energy, the node command weakens, but the border between nodes influences remains nearly on the mediator between the two nodes (see figs. 16a and 16b). With decreasing weight, the node command stays the same, but the border between nodes influences shifts in favor of the strongest node (see figs. 16c and 16d).

The inner dynamics of the core variables of the nodes penalty models are different. Energy decrease is more progressive than weight decrease.

Despite the two “internal” differences mentioned above, there is no observable difference at the behavioral level in our experiments. However, it might be that in different experimental settings (with more nodes and a more complex sensorimotor dynamics), they would indeed cause observable effects.



6.2 Other characteristics of habituation

In this section, we discuss briefly about the characteristics of habituation our models are not addressing explicitly and that were not tested experimentally.

6.2.1 Potentiation By definition of potentiation, repeated blocks of habituation training and spontaneous recovery lead to progressively deeper (the effects of habituation are more pronounced) and more rapid habituation (Rankin et al. 2009).

With the models proposed in this paper, potentiation can be observed only when recovery is not complete. If recovery is complete, for example with a long quiet phase, there will be no potentiation with the proposed models because all the core components are back to their initial values. If, however, recovery is not complete, for example with a short quiet phase, we will observe a decreased response on a following stimulation phase, more pronounced than in previous stimulation phases. This is explained by the fact the core components are not back to their initial values and still induce some habituation.

It is unclear to us if this corresponds to potentiation as described by biologists. With the proposed models, the “speed” of habituation is constant but starts with an already habituated response when the recovery period is short.

The proposed models cannot account for potentiation as habituation on a “faster” pace.

6.2.2 Intensity of stimulus effect Another property of habituation observed by biologists is about the effect of stimulus intensity where a weaker stimulus leads to a faster or stronger response decrement.

It is unclear to which extent our models could account for that observation. In fact, we would need to dig into more details about the setups used by the biologists to exhibit and quantify that observation.

6.2.3 Below-zero habituation Below-zero habituation is defined as that even after the response has decremented to no response or to asymptotic level, the effects of stimulation may continue to accumulate.

All the models can account for that observation. By definition (see eq. (2)), the velocity command μ can continue to accumulate after the response is completely inhibited. It can accumulate beyond the boundaries of the sensorimotor space SM , and if we did not reset the agent between phases (or clamp it to these boundaries), it will take more and more time for the behavior to recover.

6.2.4 Stimulus specificity In our setup, the agent only has one sensor. We think of it as a stimulus specific sensor, able to perceive a given stimulus of varying intensity. Thus in order to evaluate the stimulus specificity of habituation, we would need to instantiate the agent with several sensors. With the sensor memory model, if the agent has a memory for each of its sensors, then the habituation would show stimulus specificity. For the nodes penalty models, it would depend on the position of the nodes in the sensorimotor space, and the minimum distance between a node and the agent sensorimotor state for the penalty to significantly affect the node component.

6.2.5 Dishabituation Dishabituation appears as an increase of the habituated response when a novel/noxious stimulus is inserted into a series of habituation stimuli. For this to happen, in contrast to the stimulus specificity, we need some dependency or interaction between stimuli. This is not implemented in the proposed models.

6.2.6 Habituation of dishabituation Since there is no dishabituation mechanism currently in our models, there is no habituation of dishabituation. Nonetheless, if this characteristic relies on inhibition of the perception of the stimulus responsible of dishabituation, the repetition of said stimulus should indeed triggers habituation of dishabituation in the sensor memory model.

6.2.7 Long-term habituation Long term habituation which stands for response decrement lasting hours, days or weeks can be accounted by tuning the parameters of the proposed models resulting in habituation being slower to take place and slower to recover. If there is an asymmetry in the time to habituate with respect to the time to recover, for the penalty models, habituation and recovery can be independently tuned and for the sensor memory model, both habituation and recovery takes place on the same time scale. We could also combine multiple models.

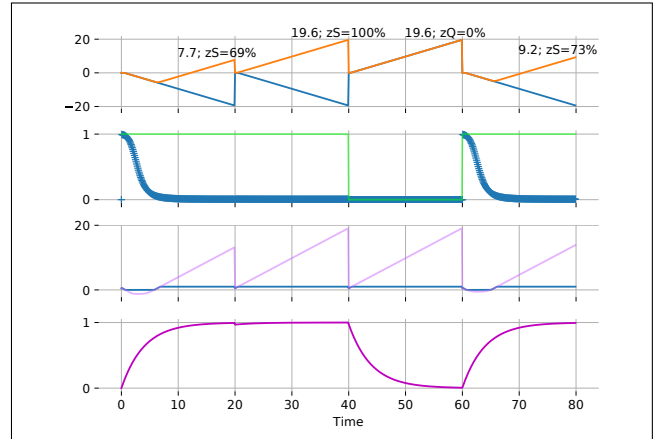


Figure 17. Sensor memory mechanism with the original IDSM. From top to bottom: the trajectory of the learning and control agents, the stimulus and agent's sensor values, the normalized motor value and velocity command μ , the agent sensor memory δ . Habituation is still observed even if new nodes are created and their weights continuously adapted.

6.3 Impact on full “deformable” IDSM

In our models, we make the strong hypothesis that the weight of nodes are kept constant and that no new nodes are created, hence our work can be rather considered as being based on a simplified variant of IDSM called “node based sensorimotor map” (NB-SMM) Woolford and Egbert (2020). These two mechanisms have been dropped in our study to simplify the analysis of our models.

However, these two mechanisms are critical for the deformable part of the IDSM which allow for the encoding and maintenance of habits. And it is thus important to discuss in more details the compatibility of our models with the full IDSM. In very preliminary works with the sensor memory and node energy mechanisms with both the creation of nodes and the modification of the weights as in the original IDSM, we still observe habituation.

Figure 17 illustrates the application of the sensory memory model to a full IDSM, starting with 2 nodes as in fig. 4 and we still observe habituation. A more in depth analysis is required but our hypothesis of why it works is the following. In addition to the first two nodes, new nodes are created as the sensorimotor state of the agent moves in the sensorimotor space. Even if new nodes are created, the sensory memory progressively “flattens” the sensorimotor trajectory to lie in a region close to the node N_2 responsible for the default upward motion. At that point, no new nodes are created, nodes pushing the agent up are reinforced and other, older, nodes are weakened, thus reinforcing the habituation effect. In the recovery phase, old nodes still exists and are reactivated, leading to the previous downward response behavior.

Figure 18 depicts preliminary work on using the energy model to a full IDSM, starting again with 2 nodes as in fig. 4. In the first subplot (a) the energy mechanisms work as expected on the energy levels of nodes N_1 and N_2 , but the resulting behavior is very close to the control behavior, thus not exhibiting an habituation behavior. The

main reason is that the motor command μ builds such inertia for going up that habituation is not even needed for the upward motion to be expressed. In the second plot (b), when we clamp μ , the energy levels allow the expression of an habituation effect. It seems that the energy mechanism is compatible with the full IDSM but we must not forget that, in our experiments with the energy model alone, the sensor component is either 0 or 1, with no intermediary values. Still, another preliminary experiment where we mix sensory memory and energy model, the full IDSM has to deal with sensorimotor trajectories where the sensor component varies continuously in $[0, 1]$. As shown on fig. 19, our combined model seem to exhibit habituation. But, again, more experiments and in depth analysis should be conducted to assess the compatibility of sensory memory and energy mechanisms with IDSM, compatibility which must not be taken as a given.

The case of the weight penalty mechanism is harder to mix with a full IDSM model as weight decay and weight penalties should be very carefully parameterized for deformant habit formation and habituation to appear. None of our preliminary experiments were conclusive.

7 Conclusion

In this paper, we proposed several extensions of the IDSM to account for some of the multiple facets of habituation. These models focused on the two main properties of habituation: inhibition of the response and spontaneous recovery. They exhibit different characteristics and combining them provide better models. The proposed models could be considered as a basis, which could be refined, combined or extended for accounting for all the properties of habituation as defined by the biologists.

Having an operational model of habituation, a fundamental and widespread mechanism, also allows to question in a practical way how habituation contributes to the adaptation of an organism to its environment. That question is not only interesting from a biological perspectives but also for artificial agents. How can habituation help artificial agents acquire new skills?

References

- Barandiaran XE and Di Paolo EA (2014) A genealogical map of the concept of habit. *Frontiers in Human Neuroscience* 8: 522.
- Bi S, Diao Q, Chai X and Han C (2017) On a neural network model based on non-associative learning mechanism and its application. In: *2017 6th Data Driven Control and Learning Systems (DDCLS)*. IEEE, pp. 679–684.
- Egbert M (2018) Investigations of an adaptive and autonomous sensorimotor individual. In: *ALIFE 2018: The 2018 Conference on Artificial Life*. MIT Press, pp. 343–350.
- Egbert MD and Barandiaran XE (2014) Modeling habits as self-sustaining patterns of sensorimotor behavior. *Frontiers in human neuroscience* 8: 590.
- Eisenstein E, Eisenstein D and Smith JC (2001) The evolutionary significance of habituation and sensitization across phylogeny: A behavioral homeostasis model. *Integrative Physiological & Behavioral Science* 36(4): 251–265.
- Hall G and Rodríguez G (2020) When the stimulus is predicted and what the stimulus predicts: Alternative accounts of habituation. *Journal of Experimental Psychology: Animal Learning and Cognition* 46(3): 327.
- Hong Q, Yan R, Wang C and Sun J (2020) Memristive circuit implementation of biological nonassociative learning mechanism and its applications. *IEEE Transactions on Biomedical Circuits and Systems* 14(5): 1036–1050.
- Marsland S (2009) Using habituation in machine learning. *Neurobiology of learning and memory* 92(2): 260–266.
- Marsland S, Nehmzow U and Shapiro J (1999) A model of habituation applied to mobile robots. *Proceedings of towards intelligent mobile robots*.
- Marsland S, Nehmzow U, Shapiro J et al. (2000) Detecting novel features of an environment using habituation. In: *Proc. Simulation of Adaptive Behavior*. Citeseer.
- McDiarmid TA, Yu AJ and Rankin CH (2019) Habituation is more than learning to ignore: multiple mechanisms serve to facilitate shifts in behavioral strategy. *BioEssays* 41(9): 1900077.
- Randlett O, Haesemeyer M, Forkin G, Shoenhard H, Schier AF, Engert F and Granato M (2019) Distributed plasticity drives visual habituation learning in larval zebrafish. *Current Biology* 29(8): 1337–1345.
- Rankin CH, Abrams T, Barry RJ, Bhatnagar S, Clayton DF, Colombo J, Coppola G, Geyer MA, Glanzman DL, Marsland S et al. (2009) Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiology of learning and memory* 92(2): 135–138.
- Rankin CH, Beck CD and Chiba CM (1990) *Caenorhabditis elegans*: a new model system for the study of learning and memory. *Behavioural brain research* 37(1): 89–92.
- Stanley JC (1976) Computer simulation of a model of habituation. *Nature* 261(5556): 146–148.
- Thompson RF (2009) Habituation: a history. *Neurobiology of learning and memory* 92(2): 127.
- Thompson RF and Spencer WA (1966) Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychological review* 73(1): 16.
- Vogel D and Dussutour A (2016) Direct transfer of learned behaviour via cell fusion in non-neural organisms. *Proceedings of the Royal Society B: Biological Sciences* 283(1845): 20162382.
- Vogel EH, Ponce FP and Wagner AR (2019) The development and present status of the sop model of associative learning. *Quarterly Journal of Experimental Psychology* 72(2): 346–374.
- Wagner A (1981) Sop: A model of automatic memory processing in animal behavior. *Information processing in animals: Memory mechanisms*.
- Wang D (1993) A neural model of synaptic plasticity underlying short-term and long-term habituation. *Adaptive Behavior* 2(2): 111–129.
- Wang D and Hsu C (1990) Slonn: A simulation language for modeling of neural networks. *Simulation* 55(2): 69–83.
- Woolford FM and Egbert MD (2020) Behavioural variety of a node-based sensorimotor-to-motor map. *Adaptive Behavior* 28(6): 425–440.

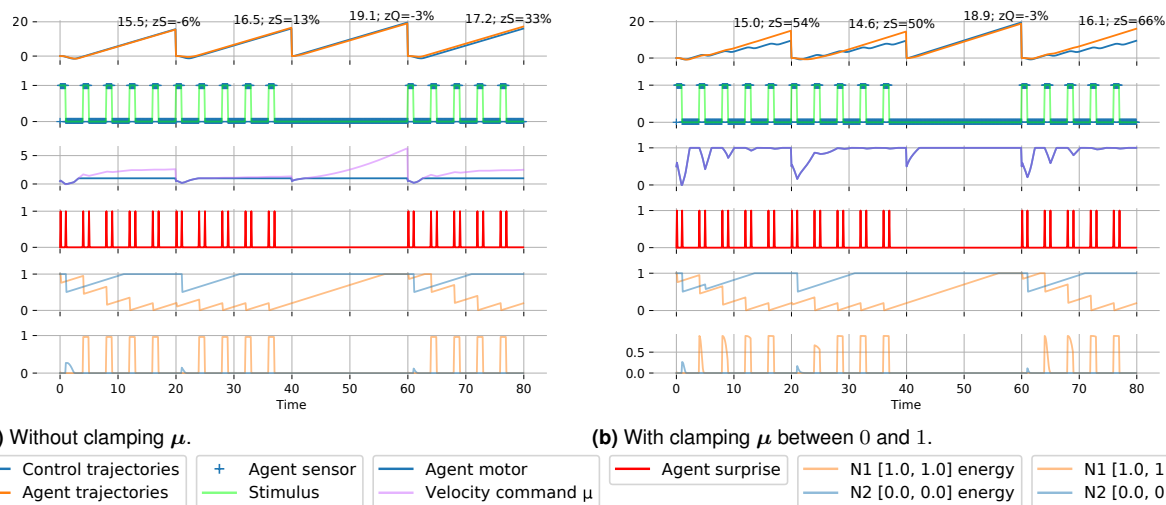


Figure 18. Node Energy mechanism with the original IDSM. From top to bottom: the trajectory of the learning and control agents, the stimulus and agent's sensor values, the normalized motor value and velocity command μ , the surprises, the nodes energy and the nodes ξ_N . Nodes energy level behave “as expected”, but the habituation behavior cannot express when μ is not clamped (plot (a)), but is apparent when μ is clamped (plot (b)).

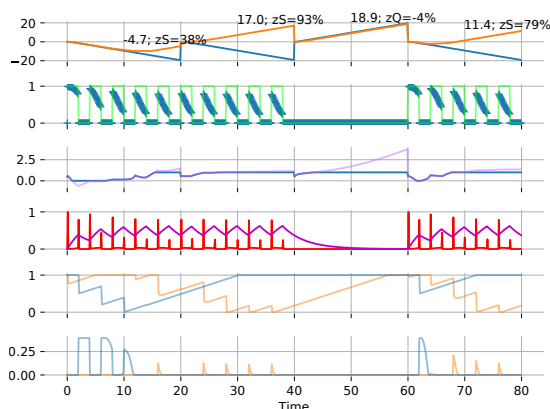


Figure 19. Sensor memory mechanism and nodes energy penalty mechanism with the original IDSM. From top to bottom: the trajectory of the learning and control agents, the stimulus and agent's sensor values, the normalized motor value and velocity command μ , the agent sensor memory δ and the surprises, the nodes energy and the nodes ξ_N . Habituation is still observed with new nodes are created, weights continuously adapted and combination of two models.

Appendix

Symbol	Short description
x_{SM}	Current SM state of the agent
x_E	Current state of the agent in the environment
S_E	Current state of the environment
x_c	Last position reached by the control agent in the absence of stimulation
x_c^*	Highest last position reached by the control agent in the absence of stimulation
N_p	SM-position associated with node N
N_v	SM-velocity indicated by node N
N_w	Scalar, weight of Node N
$\phi_N(x_{SM})$	Density of Nodes at current agent SM-position
$\omega(N_w)$	Function describing how the weight of a node scales its influence
$d(x_{SM}, y_{SM})$	Distance between two SM-positions
$V(x_{SM})$	Velocity factor (i.e. how nodes impact the current SM-velocity change)
$A(x_{SM})$	Attraction factor (i.e. the attraction factor of influent nodes on the current SM-position)
μ	Motor command of the agent, influenced by $V(x_{SM})$ and $A(x_{SM})$
ξ	Handful notation for the product $\omega(N_w)d(x_{SM}, y_{SM})$
nhs	Non habituated sensor value
hs	Habituated sensor value
δ	Sensor memory
sur	Surprise
$c_e, c_w, \gamma_w, \beta_w, \beta_e, \gamma_e$	Parameters of the energy and node weight penalty mechanisms
z_S	Score for the stimulation phase, measures the efficiency of habituation and spontaneous recovery
z_Q	Score for the quiet phase, it measures the lasting effect of reverse habituation
dt	Euler integration step (usually 0.01 time unit)
k_w	Weight parameter (by default 0.025)
k_d	Distance parameter (by default 30)

Author Biographies



Tristan Gillard is a PhD student in Computer Science at University de Lorraine. He works on unconventional computing approaches inspired by biology in the Biscuit team of the Loria laboratory, in particular on unsupervised artificial agent.



Jeremy Fix is an assistant professor in computer science at CentraleSupélec. After completing a PhD in Computer Science from the Université Henri Poincaré on computational models of anticipation in 2008, a postdoctorate at the Münster University in psychology working on visual attention models from in vivo recordings, a postdoctorate at Supelec on distributed computational models of cortical functions, Jeremy Fix is since 2013 an assistant professor at CentraleSupélec. He participates to the teaching of machine learning at CentraleSupélec, especially deep learning and carries research on unconventional computing approaches in the Biscuit team of the Loria laboratory. He also supervises PhD students on machine learning applied on topics such as inner wood defect prediction and biodiversity estimation.



Alain Dutech is a research fellow in artificial intelligence at INRIA. He carries research on unconventional computing approaches in the Biscuit team of the Loria research lab in Nancy, France. His topics of research encompass reinforcement learning, unsupervised learning and self-organizing systems.