



**HAL**  
open science

# Non-Associative Learning and the Iterant Deformable Sensorimotor Medium

Tristan Gillard, Alain Dutech, Jérémy Fix

► **To cite this version:**

Tristan Gillard, Alain Dutech, Jérémy Fix. Non-Associative Learning and the Iterant Deformable Sensorimotor Medium. [Research Report] RR-9472, Inria Nancy - Grand Est. 2022, pp.50. hal-03672726

**HAL Id: hal-03672726**

**<https://inria.hal.science/hal-03672726>**

Submitted on 20 May 2022

**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.

*Inria*

# Non-Associative Learning and the Iterant Deformable Sensorimotor Medium

Tristan GILLARD, Jérémy FIX, Alain DUTECH

**RESEARCH  
REPORT**

**N° 9472**

May 2022

Project-Teams BISCUIT

ISRN INRIA/RR--9472--FR+ENG

ISSN 0249-6399





## Non-Associative Learning and the Iterant Deformable Sensorimotor Medium

Tristan GILLARD<sup>\*†</sup>, Jérémy FIX<sup>‡†</sup>, Alain DUTECH<sup>§ †</sup>

Project-Teams BISCUIT

Research Report n° 9472 — May 2022 — 50 pages

**Abstract:** What drives the behavior of an organism ? That fundamental question has been addressed in the literature for a variety of organisms and through different angles. For some species, part of the answer lies in the associative learning mechanisms but there also exists more widespread non-associative learning mechanisms. *Habituation* and *Sensitization*, the first translating behaviorally in a response decrease while the second in a response increase, are two such mechanisms that are so widely observed in the phylogeny that it suggests their fundamental role in behavioral learning.

In this context, we propose new computational mechanisms of learning inspired by habituation and sensitization as described by biologists. These models are seen as extensions to the Iterant Deformable Sensorimotor Medium (IDSM), 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.

**Key-words:** autonomous agent, non-associative learning, habituation, sensitization, behavior formation

---

\* Loria - Univ. Lorraine

† BISCUIT Team

‡ Loria - CentraleSupélec, Campus de Metz

§ Loria - INRIA

**RESEARCH CENTRE  
NANCY – GRAND EST**

615 rue du Jardin Botanique  
CS20101  
54603 Villers-lès-Nancy Cedex

## Apprentissage Non-Associatif et le modèle “Iterant Deformable Sensorimotor Medium”

**Résumé :** Qu’est-ce qui motive le comportement d’un organisme ? Cette question fondamentale a été abordée dans la littérature pour une variété d’organismes et sous différents angles. Pour certaines espèces, une partie de la réponse réside dans les mécanismes d’apprentissage associatif, mais il existe également des mécanismes d’apprentissage non associatif plus répandus. *Habituation* et *Sensibilisation*, le premier se traduisant comportementalement par une diminution de la réponse et le second par une augmentation de la réponse, sont deux de ces mécanismes qui sont si largement observés dans la phylogénie que cela suggère leur rôle fondamental dans l’apprentissage comportemental.

Dans ce contexte, nous proposons de nouveaux mécanismes computationnels d’apprentissage inspirés de l’habituation et de la sensibilisation telles que décrits par les biologistes. Ces modèles sont considérés comme des extensions du “*Iterant Deformable Sensorimotor Medium*” (IDSM), un modèle conceptuel de la formation d’habitudes pour un agent artificiel en tant que schémas comportementaux sensorimoteurs auto-renforcés. Il s’agit d’un modèle intermédiaire entre les modèles neuronaux et les modèles de comportement macroscopiques.

**Mots-clés :** agent autonome, apprentissage non-associatif, habituation, sensibilisation, formation de comportements

# Contents

<b>1</b>	<b>IDSM</b>	<b>7</b>
1.1	Formal description of the IDSM . . . . .	7
1.2	Limitations of IDSM . . . . .	9
<b>2</b>	<b>Habituation Models</b>	<b>11</b>
2.1	Introduction . . . . .	11
2.2	Habituation . . . . .	12
2.2.1	The characteristics of habituation . . . . .	12
2.2.2	Experimental IDSM setup for Habituation . . . . .	14
2.2.3	Expected behaviors . . . . .	14
2.3	Models of Habituation . . . . .	15
2.3.1	Habituation through sensor memory . . . . .	15
2.3.2	Habituation through nodes penalty . . . . .	16
2.4	Experiments and results . . . . .	18
2.4.1	Habituation and spontaneous recovery experimental setup . . . . .	18
2.4.2	Habituation and spontaneous recovery results . . . . .	18
2.4.3	Interstimulus interval setup and results . . . . .	23
2.4.4	Combination of models . . . . .	28
2.5	Discussion about Habituation . . . . .	28
2.5.1	Models comparison . . . . .	28
2.5.2	Other characteristics of habituation . . . . .	33
<b>3</b>	<b>Sensitization Models</b>	<b>37</b>
3.1	Introduction . . . . .	37
3.2	Sensitization . . . . .	38
3.3	Models . . . . .	39
3.3.1	Site-specific sensitization . . . . .	39
3.3.2	Pseudo-conditioning through generalization of nodes . . . . .	41
3.4	Experiments and results . . . . .	42
3.4.1	Site-specific sensitization setup . . . . .	42
3.4.2	Site-specific sensitization results . . . . .	43
3.4.3	Pseudo-conditioning setup . . . . .	44
3.4.4	Pseudo-conditioning results . . . . .	45
3.5	Discussion about Sensitization . . . . .	46



# General Introduction

This technical report presents the various computational models of two non-associative learning mechanisms (Habituation, see chapter 2 and Sensitization, see chapter 3) that have been developed and studied during the PhD work of Tristan GILLARD. All these models are thought as extensions to the Iterant Deformable Sensorimotor Medium (IDSM, see chapter 1) defined by (Egbert and Barandiaran, 2014), a high level computational model of habits formation.

Each of the presented models is implemented and experimented on archetypal scenarii so has to present its characteristics and limits.

The main objective of the PhD work of Tristan GILLARD is to explore new unsupervised learning mechanisms that could expand the behavior skills of an artificial agent.



# Chapter 1

## IDSM

### 1.1 Formal description of the IDSM

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 1.1, part of the environment state  $\mathbf{S}_E$  is abstracted into a sensorimotor state  $\mathbf{S}_{SM}$ . In the sensorimotor space, the IDSM uses  $\mathbf{S}_{SM}$  and some *nodes* that store preferred sensorimotor alterations to compute a command  $\boldsymbol{\mu}$ . This command  $\boldsymbol{\mu}$ , 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:

$$N = \langle \mathbf{p}, \mathbf{v}, w \rangle \quad (1.1)$$

with

- $\mathbf{p}$  a vector indicating the node's position in sensorimotor space SM,
- $\mathbf{v}$  a vector indicating the desired velocity in sensorimotor space SM at the position  $\mathbf{p}$ ,
- $w$  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 1.2 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  $N_{\mathbf{p}}$  and the velocity component seeking to align the agent's velocity with  $N_{\mathbf{v}}$ . 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 (1.2)-(1.7)). The velocity command  $\boldsymbol{\mu}$  of the agent, normalized in the sensorimotor space SM, is given by:

$$\frac{d\boldsymbol{\mu}}{dt} = \begin{cases} \frac{V(\mathbf{x}_{SM}) + A(\mathbf{x}_{SM})}{\phi_N(\mathbf{x}_{SM})} & \text{if } \phi_N(\mathbf{x}_{SM}) > \epsilon \\ 0 & \text{otherwise} \end{cases} \quad (1.2)$$

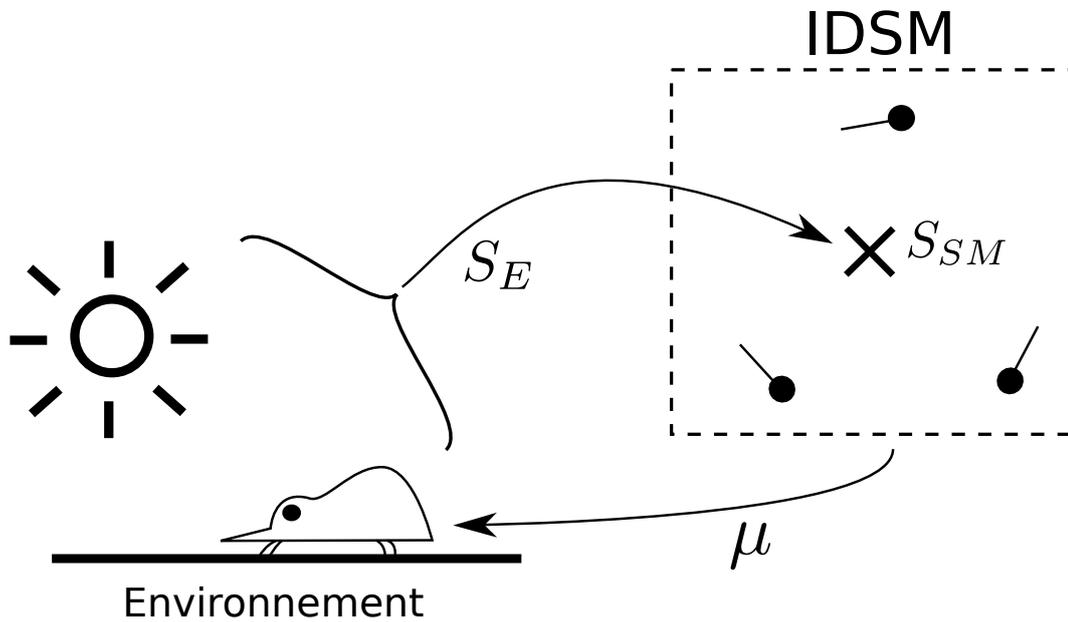


Figure 1.1: An IDSM works in the sensorimotor space of the agent. According to the current sensorimotor state  $S_{SM}$  and to the various nodes of the IDSM, it produces commands  $\mu$  that control the agent behavior in its environment.  $S_E$  is the complete state of the environment.

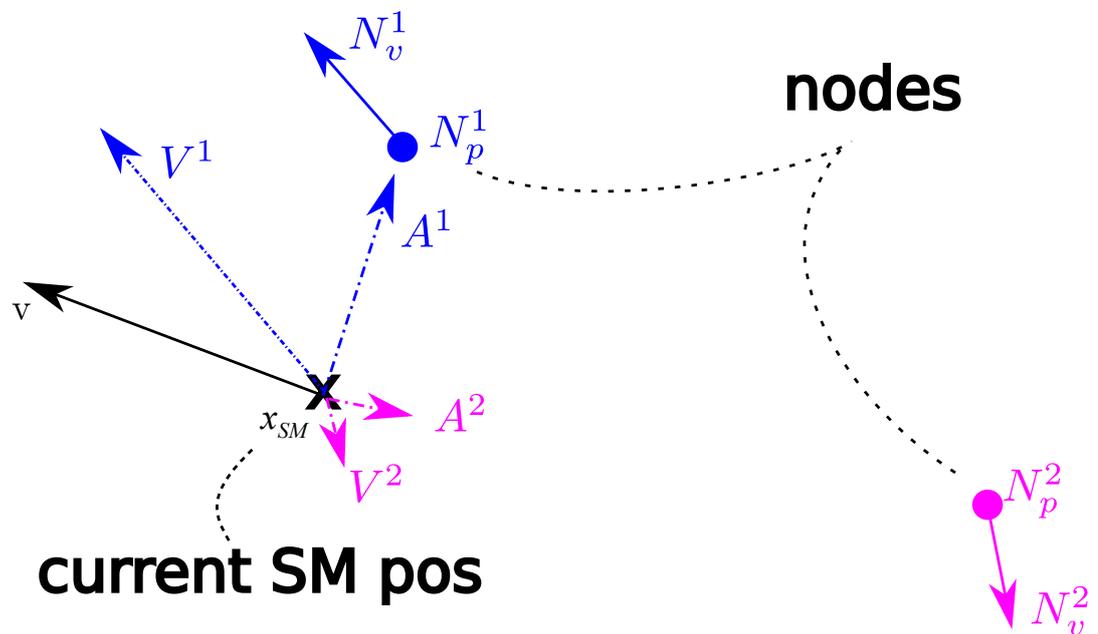


Figure 1.2: An IDSM is built from a collection of nodes  $N^i$  which influence the agent dynamics with a velocity  $V^i$  and attraction  $A^i$  components.

with the velocity  $V$  and attraction  $A$  components defined by equations (1.3)-(1.8).

$$V(\mathbf{x}_{SM}) = \sum_{\mathcal{N}} \omega d_N \cdot [N_{\mathbf{v}}]^\mu \quad (1.3)$$

$$A(\mathbf{x}_{SM}) = \sum_{\mathcal{N}} \omega d_N \cdot [\Gamma(N_{\mathbf{p}} - \mathbf{x}_{SM}, N_{\mathbf{v}})]^\mu, \quad (1.4)$$

$$\omega d_N = \omega(N_w) \cdot d(N_{\mathbf{p}}, \mathbf{x}_{SM}) \quad (1.5)$$

$$\omega(N_w) = \frac{2}{1 + \exp(-k_\omega N_w)} \quad (1.6)$$

$$d(N_{\mathbf{p}}, \mathbf{x}_{SM}) = \frac{2}{1 + \exp(k_d \|N_{\mathbf{p}} - \mathbf{x}_{SM}\|^2)} \quad (1.7)$$

$$\Gamma(\mathbf{a}, N_{\mathbf{v}}) = \begin{cases} \mathbf{a} - \left\langle \mathbf{a}, \frac{N_{\mathbf{v}}}{\|N_{\mathbf{v}}\|} \right\rangle \frac{N_{\mathbf{v}}}{\|N_{\mathbf{v}}\|} & \text{if } N_{\mathbf{v}} \neq 0 \\ \mathbf{a} & \text{otherwise} \end{cases} \quad (1.8)$$

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

In equation (1.2), the local density of the nodes  $\phi_{\mathcal{N}}(\mathbf{x}_{SM})$  is computed using:

$$\phi_{\mathcal{N}}(\mathbf{x}_{SM}) = \sum_{\mathcal{N}} \omega d_N \quad (1.9)$$

The differential equation (1.2) 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.

- Alain]

## 1.2 Limitations of IDSM

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. 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.



## Chapter 2

# Habituation Models

### 2.1 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 2.2. 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. [DEL: In a future work, we will explore how these models can help artificial agents acquire and develop new behaviors. - **Alain**]

Habituation is a behavioral property which has been studied for a quite a long time, especially since the seminal works of Rescorla-Wagner which 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 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 chapter about habituation, 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 (IDS<sub>M</sub>) (Egbert and Barandiaran, 2014) described in section 1.1. Our motivation is to add learning or adaptation mechanisms to the IDS<sub>M</sub> framework so as to open the possibility for the artificial agent to produce new behaviors.

[MOV: We are interested in artificial agents that can acquire new behaviors and the IDS<sub>M</sub> framework specifically addresses the problem of using a simple kind of basic behaviors: habits. IDS<sub>M</sub> 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. - Alain]

In this context, we propose two family of extensions to the IDS<sub>M</sub> framework inspired by habituation. One is based on the idea of having a *sensor memory* which impacts the response triggered by a stimulus. 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. This paper is structured as follows. We start by recalling the biological definition of habituation and its behavioral correlates in section 2.2. The Iterant Deformable Sensorimotor Medium (IDS<sub>M</sub>) framework is then described in section 1.1. The two proposed family of models are detailed in section 2.3 and various experiments are presented and analyzed in section 2.4. We conclude the paper with a discussion on our models. The code used for the experiments of this paper is available online<sup>1</sup>.

## 2.2 Habituation

### 2.2.1 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.

<sup>1</sup>[https://gitlab.inria.fr/openbiscuit/papercode/2021\\_habituationgillard](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 2.1: Characteristics of habituation from (McDiarmid et al., 2019; Rankin et al., 2009)

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 2.1. These characteristics are discussed in light of our models in section 2.5.

[Even though not directly addressed in this work, these characteristics are discussed in light of our models in section - Alain]

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

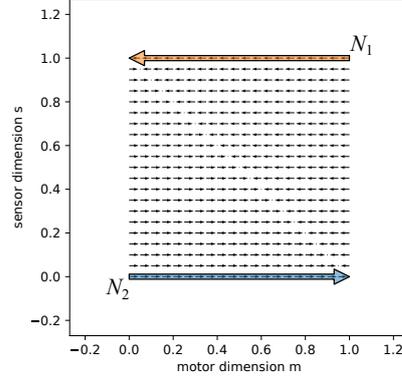


Figure 2.1: For our experiment, we consider an IDSM with two nodes. See text for details.

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.

## 2.2.2 Experimental IDSM setup for Habituation

[Originally within the IDSM description for habituation paper - Alain]

For our experiments, we build an IDSM with two nodes as shown on Figure 2.1. The two nodes are defined by:

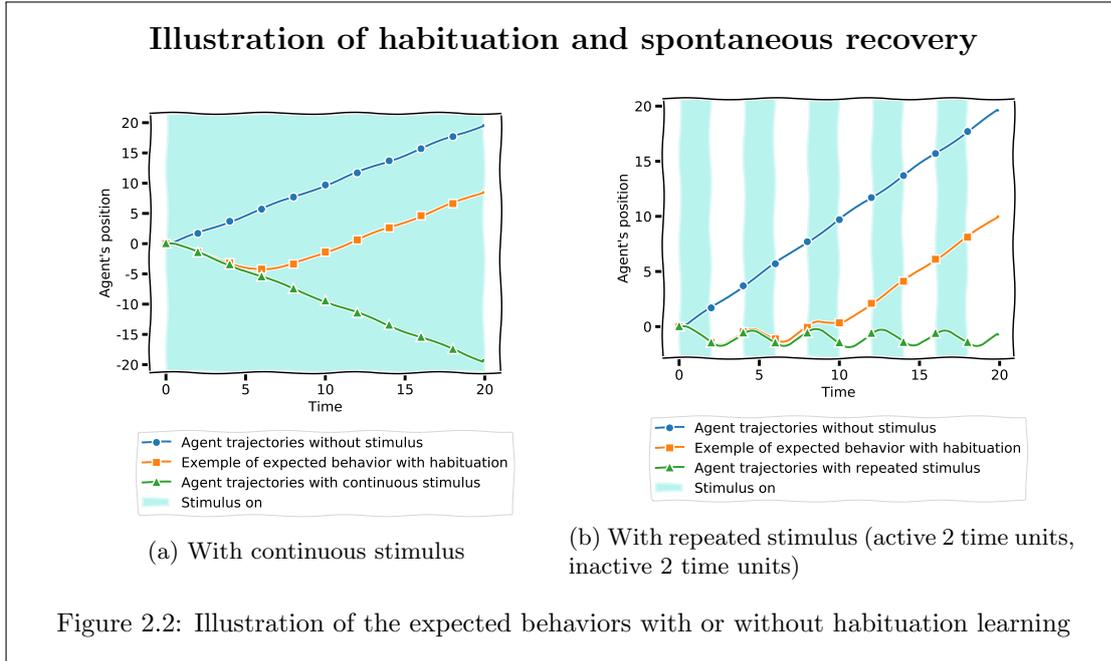
- $N_1 = \langle \mathbf{p} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, w = 0 \rangle$ ,
- $N_2 = \langle \mathbf{p} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, w = 0 \rangle$ .

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.

## 2.2.3 Expected behaviors

In this part, we illustrate the behavior we expect from our models inspired by habituation using a thought experiment, illustrated on figure 2.2. 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 2.2, 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 ease readability, the periods when the stimulus is active are represented with a colored background.

In the first setup (Figure 2.2a), the stimulus is continuously on. In the second setup (Figure 2.2b), 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.

## 2.3 Models of Habituation

In order to integrate habituation as defined in section 2.2 into the IDSM, we propose three models. In the first one (see section 2.3.1), 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 2.3.2) or through weight modification (see section 2.3.2).

### 2.3.1 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  $\delta$ , integrating the value of the sensor, and decreasing the

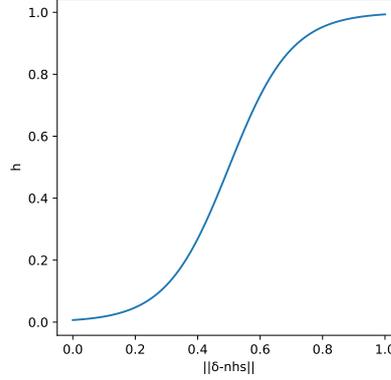


Figure 2.3: The habituation factor  $h$  as a function of the difference between the sensor value  $nhs(\mathbf{x}_E)$  and the sensor memory  $\delta$ .

agent's perception of the stimulus. This decrease in perception ultimately causes a decrease in the expression of the response. The sensor memory  $\delta$  keeps track of the evolution of the perceived stimulus. In this model, the sensor value seen by the IDSM is the *habituated* sensor value, which is modulated by a function (habituation factor  $h$ ) of the difference between the sensor value and its mean over a recent time window  $\delta$ . The higher the sensor memory  $\delta$  is, the more the motor response is inhibited, without altering the nodes in the IDSM (habituation). If the stimulus is switched off, the sensor memory  $\delta$  will decrease and then the motor response recovers progressively (spontaneous recovery).

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

$$\frac{\partial \delta}{\partial t}(\mathbf{x}_E, t) = \beta(nhs(\mathbf{x}_E) - \delta(\mathbf{x}_E, t)) \quad (2.1)$$

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  $\delta$  as:

$$hs(\mathbf{x}_E, t) = h(\mathbf{x}_E, t)nhs(\mathbf{x}_E) \quad (2.2)$$

$$h(\mathbf{x}_E, t) = \frac{1}{1 + e^a} \quad (2.3)$$

$$a = \alpha \left( \|\delta(\mathbf{x}_E, t) - nhs(\mathbf{x}_E)\|_2 - \frac{1}{2} \right) \quad (2.4)$$

with  $\alpha = -10$ . From the plot of the habituation factor  $h$  on Figure 2.3, 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 ( $\delta \approx nhs$ ).

### 2.3.2 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. 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  $\mathbf{x}_{SM}$ .

$$sur(t) = \|\mathbf{x}_{SM}(t) - \hat{\mathbf{x}}_{SM}(t)\|_2 \quad (2.5)$$

$$\hat{\mathbf{x}}_{SM}(t) = \mathbf{x}_{SM}(t - \Delta t) + \begin{bmatrix} \boldsymbol{\mu} \\ 0 \end{bmatrix} \Delta t \quad (2.6)$$

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.

### 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:

$$V(\mathbf{x}_{SM}) = \sum_{\mathcal{A}} \omega d_N \cdot N_e \cdot [N_v]^\mu \quad (2.7)$$

$$A(\mathbf{x}_{SM}) = \sum_{\mathcal{A}} \omega d_N \cdot N_e \cdot [\Gamma(N_p - \mathbf{x}_{SM}, N_v)]^\mu \quad (2.8)$$

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  $\mathbf{x}_{SM}$ .

$$c_e(N, t) = ((1 + sur(t))^{\beta_e} - 1) \cdot \omega d_N \quad (2.9)$$

$$N_e(t + \Delta t) = N_e(t) + \gamma_e \Delta t - c_e(N, t) \quad (2.10)$$

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

### 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.

$$c_w(N, t) = \beta_w \cdot sur(t) \cdot \omega d_N \quad (2.11)$$

$$N_w(t + \Delta t) = N_w(t) + \gamma_w \Delta t - c_w(N, t) \quad (2.12)$$

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

## 2.4 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.

### 2.4.1 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.

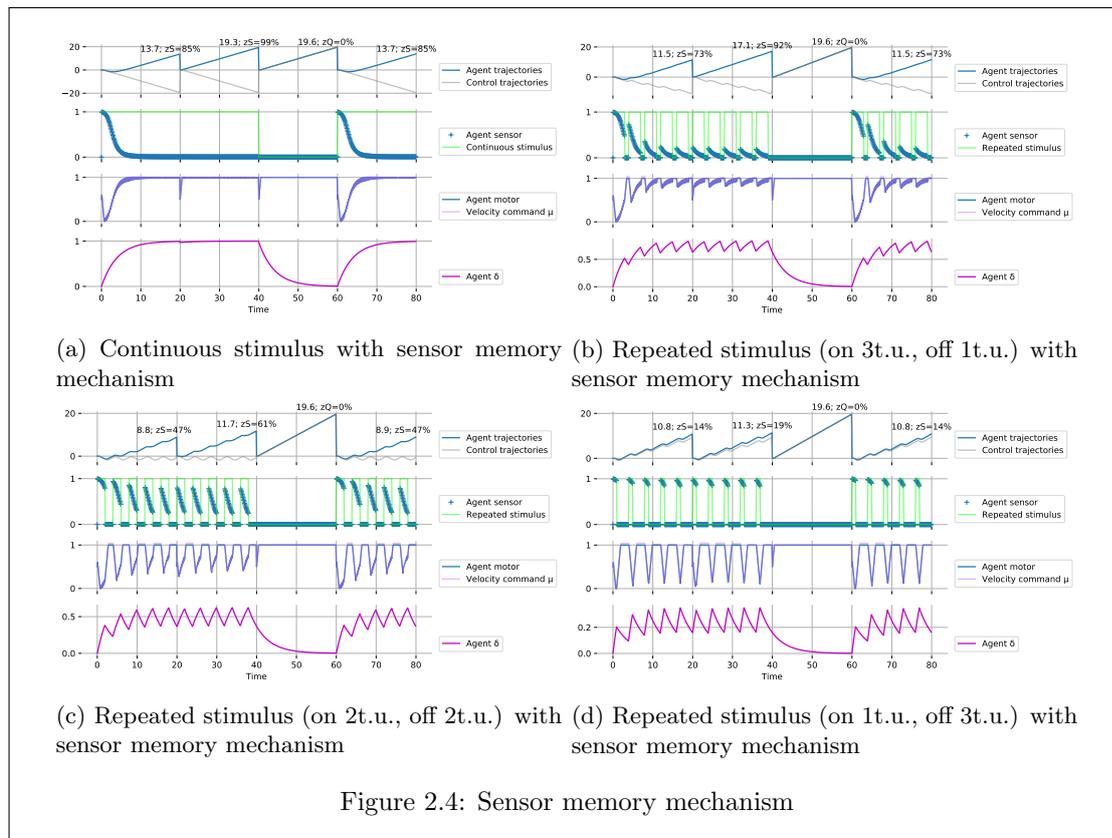
### 2.4.2 Habituation and spontaneous recovery results

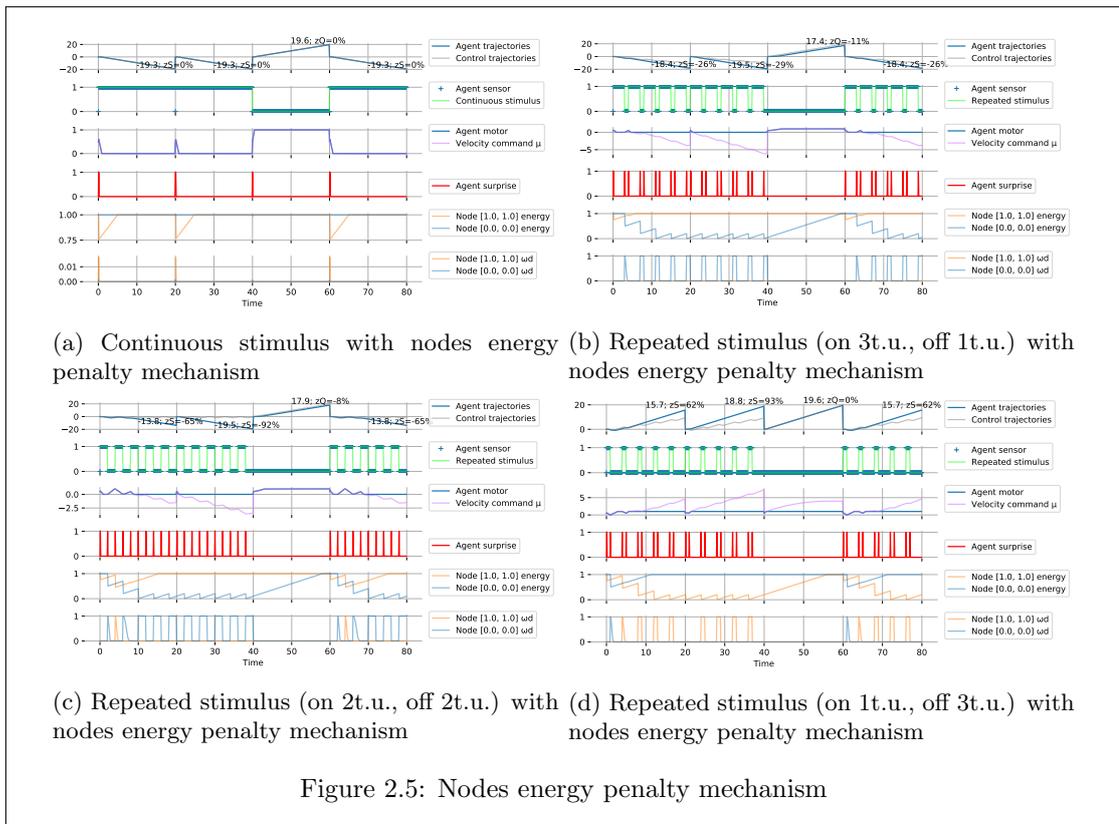
Habituation and spontaneous recovery are evaluated by comparing the behaviors of a learning agent with a baseline agent without learning mechanisms. The behaviors of the agents, for each model, are given on figures 2.4-2.6.

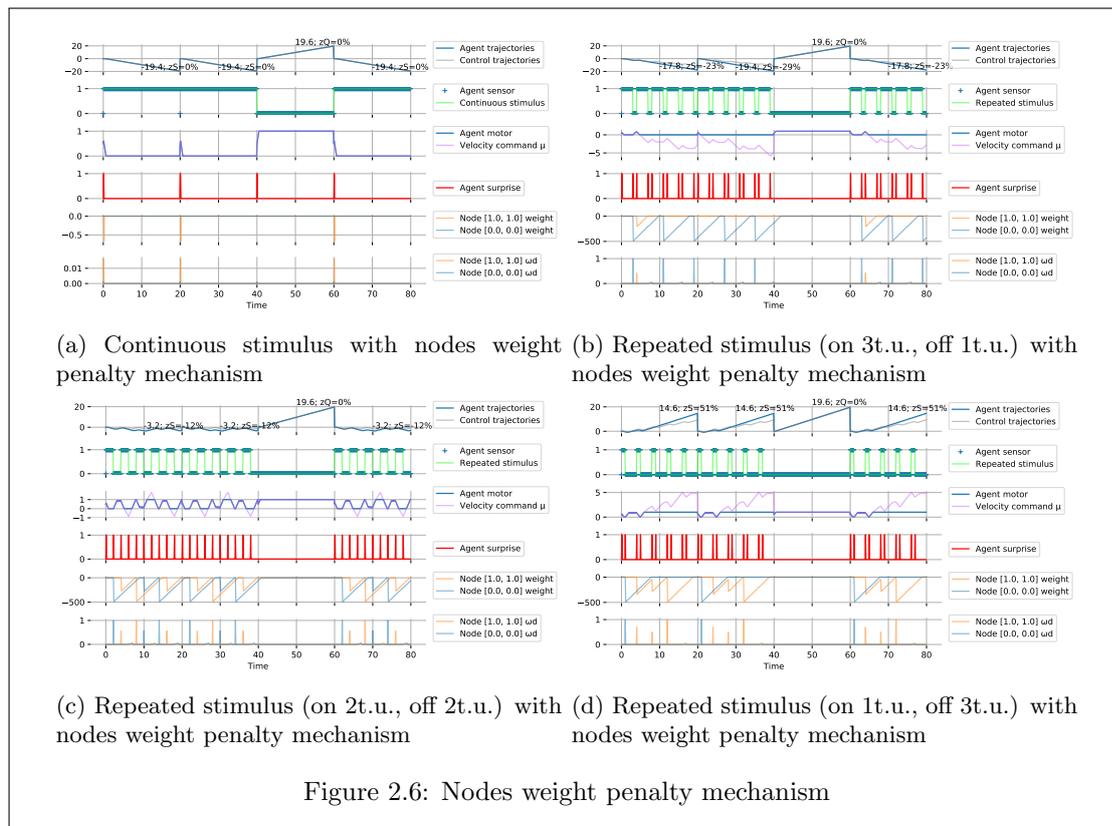
Each figure is composed of four subplots, one for each stimulation pattern. Each subplot has the same layout. The top plot displays the learning agent trajectory and the control agent trajectory. On the trajectory of the learning agent, there are two numbers: the agent last position  $\mathbf{x}_E$  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). The score for the stimulation phase measures the efficiency of habituation and spontaneous recovery, and is defined as:

$$z_S(\mathbf{x}_E) = \frac{\mathbf{x}_E - \mathbf{x}_c}{\mathbf{x}_c^* - \mathbf{x}_c} \quad (2.13)$$

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).

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

$$z_Q(\mathbf{x}_E) = -\left(1 - \frac{\mathbf{x}_E}{\mathbf{x}_c^*}\right) \quad (2.14)$$

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  $\boldsymbol{\mu}$ . The last plots give more information depending on the model (sensor memory, energy of the nodes, weights of the nodes).

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 2.3.1), as shown on figure 2.4, longer stimulus activation produces stronger habituation. This can be explained by the fact that the sensor memory  $\delta$  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 (2.3) and figure 2.3).

Conversely, nodes penalty models (sections 2.3.2 and 2.3.2) 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. 2.5 and 2.6). 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 2.7c) or low (figure 2.7d) 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 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 2.7d 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 lets 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  $\beta$ , can adjust the speed of habituation and of spontaneous recovery, but not separately. The higher  $\beta$  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  $\gamma$  (to regulate habituation) and  $\beta$  (to regulate spontaneous recovery) parameters.

### 2.4.3 Interstimulus interval setup and results

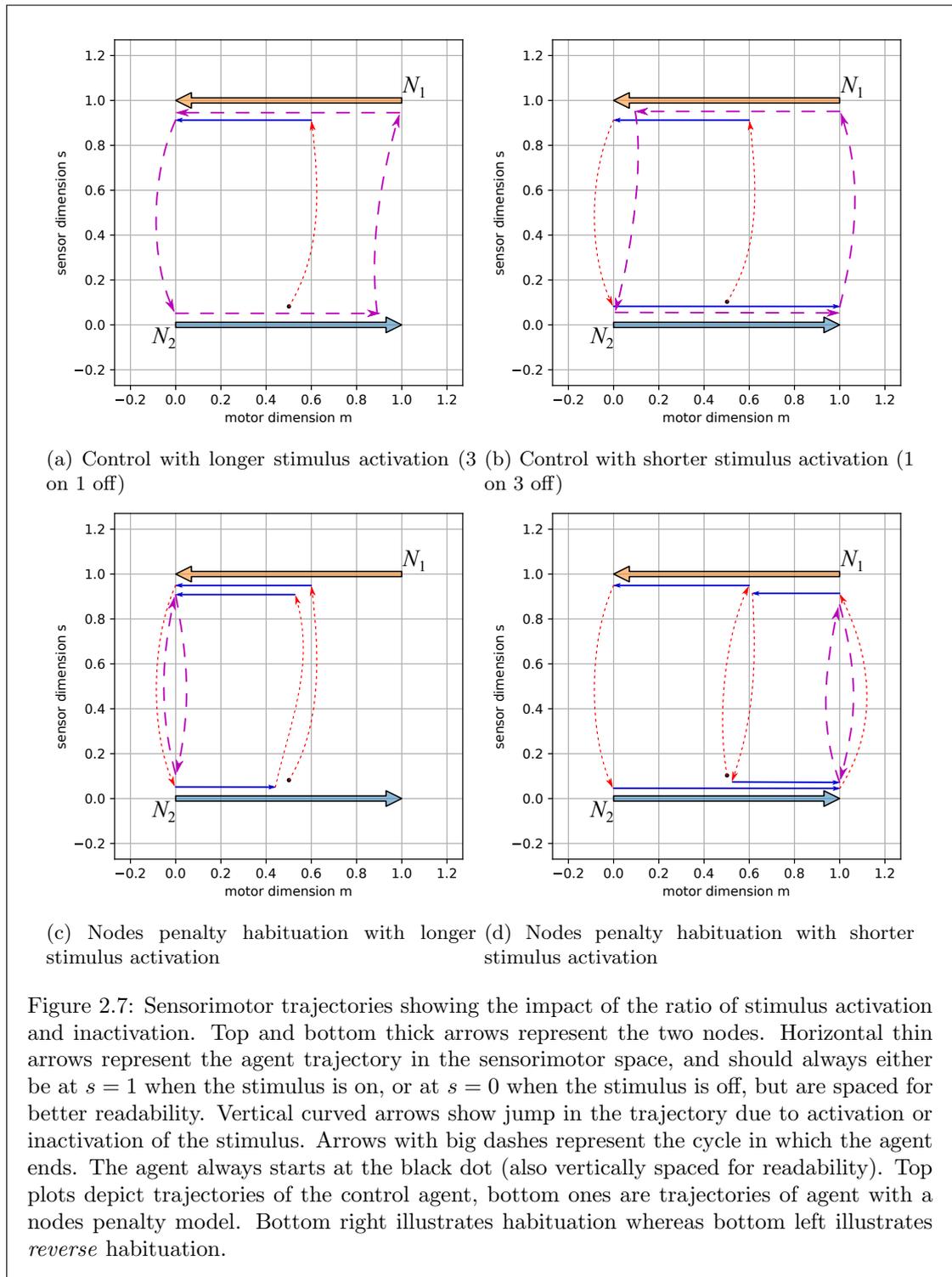
We focus here on the interstimulus interval effect which is the fourth characteristic of habituation as described in (Rankin et al., 2009). According to biologists, there are three potential consequences of higher frequency stimulation: more rapid response decrement, more pronounced response decrement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

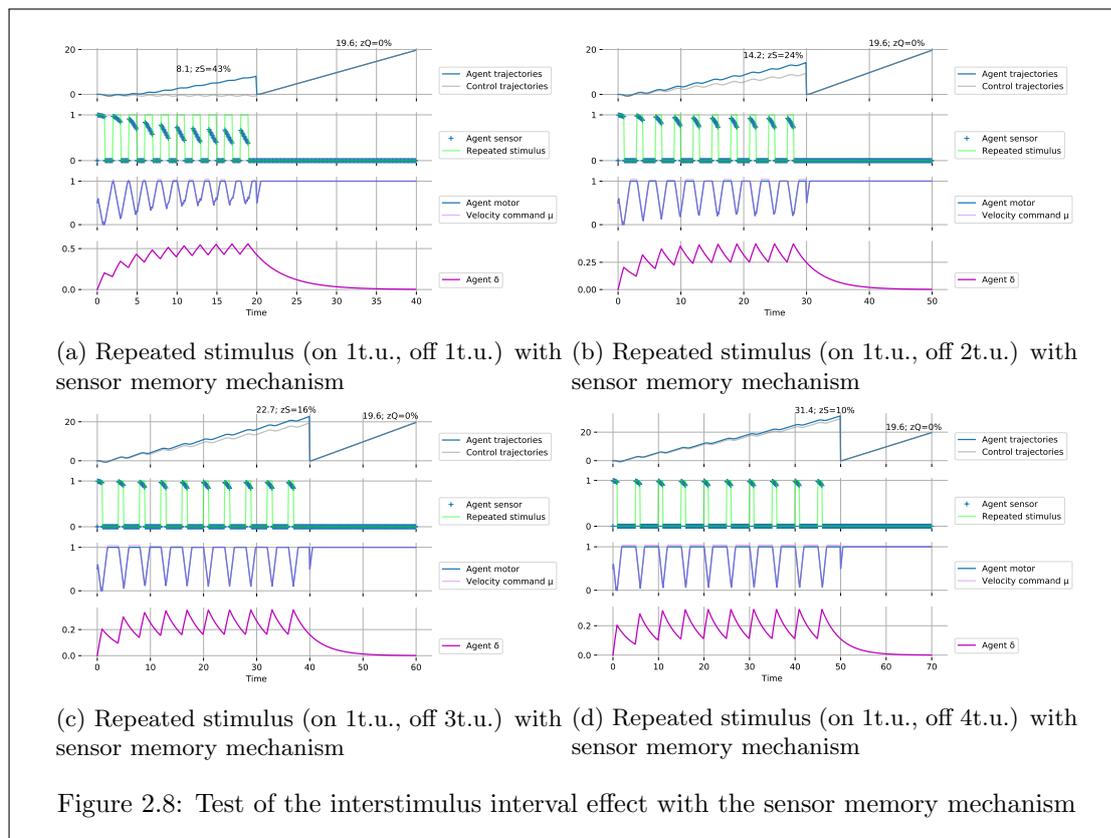
We test this characteristic by using a new protocol with two phases. During the first phase, we activate the stimulus ten times, each activation followed by an interstimulus interval, and then we have a quiet phase to evaluate the spontaneous recovery. The interstimulus interval value is set either to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

With the sensor memory model (see figure 2.8), 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  $\delta$ , 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  $\delta$  to go from respectively 0.56, 0.42, 0.35 and 0.31 to below 0.01. Thus with higher frequencies, 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:

$$r = \frac{\delta_1 - \delta_2}{t} \quad (2.15)$$

with  $\delta_1$  the value of  $\delta$  after the last deactivation of the stimulus,  $\delta_2$  the first value below 0.01 and  $t$  the time needed for the decrease between  $\delta_1$  and  $\delta_2$ . The general tendency of  $\delta$  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  $\delta$  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

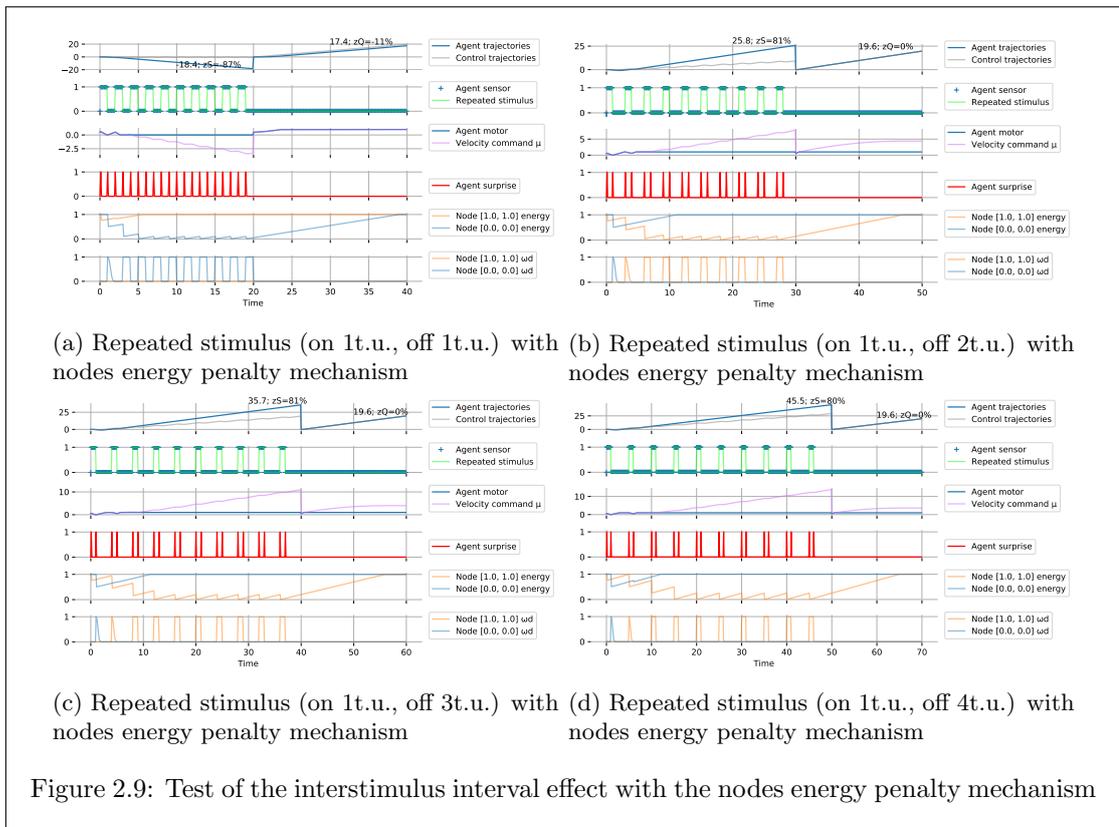


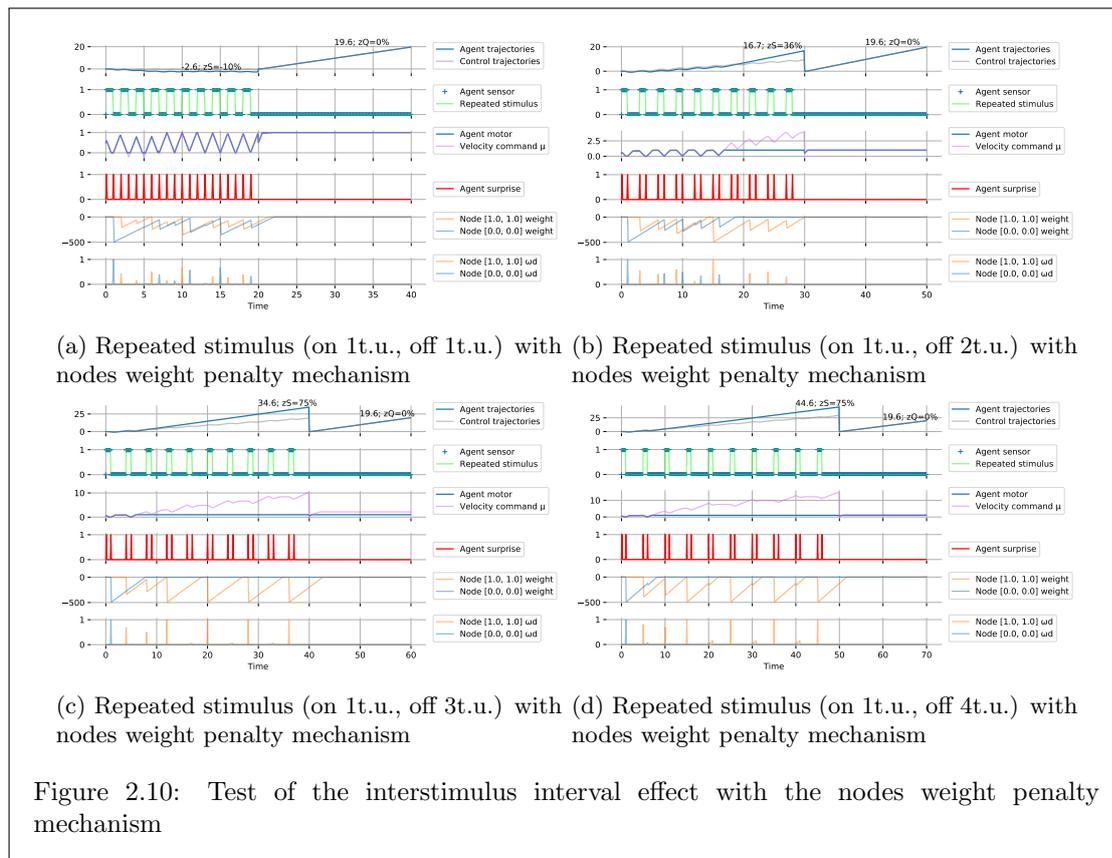


value. A crude way to evaluate the more rapid habituation effect is to measure the number of activation cycles or the time needed for  $\delta$  to reach its asymptotic mean value. In both cases, higher frequencies lead to less rapid habituation with this model: respectively, from higher to lower frequencies, 8.8, 6.7, 4.7 and 0.9 time units needed to go to the respective mean values (0.5, 0.33, 0.25 and 0.2), hence fifth, third, second and first activation cycle.

For the nodes energy penalty model (see figure 2.9), 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 2.9b, 2.9c and 2.9d, 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 2.10), 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 takes place: permanently (nodes penalty models), or directly after a the last perceived stimulus (sensor memory model).

#### 2.4.4 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 2.4.2). 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. 2.12 and 2.13). 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. 2.12 and 2.13 are always greater or equal than the habituation scores of figs. 2.4 to 2.6.

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. 2.5) and implementation of the predicted next state (eq. 2.6), and as illustrated on fig. 2.11, surprise equals the difference between two successive sensor value (represented by cross on the figure). By building up during stimulus activation, the sensor memory  $\delta$  causes a decrease in the habituated sensor value (represented by 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.

[Comming from Habituation - Alain]

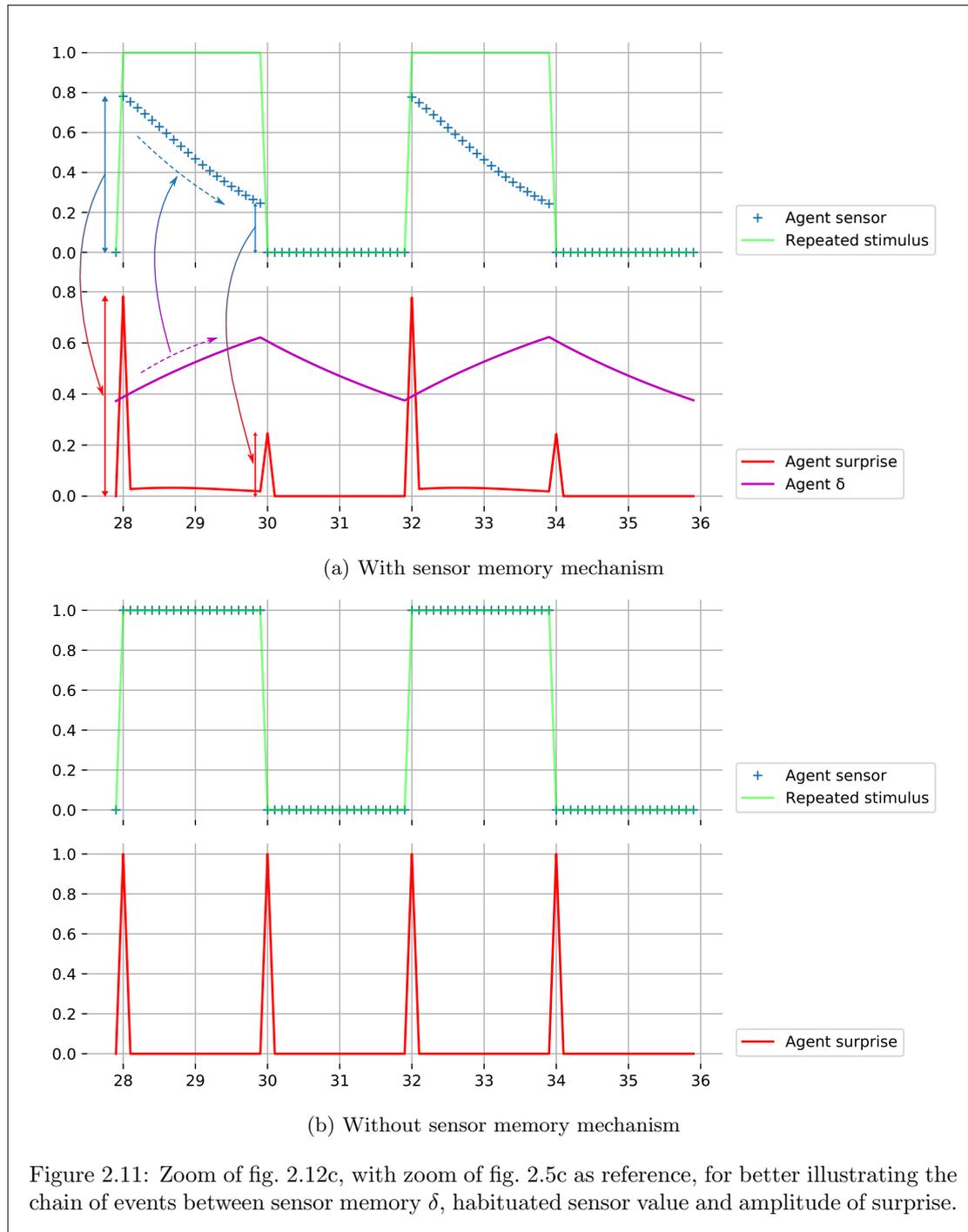
## 2.5 Discussion about Habituation

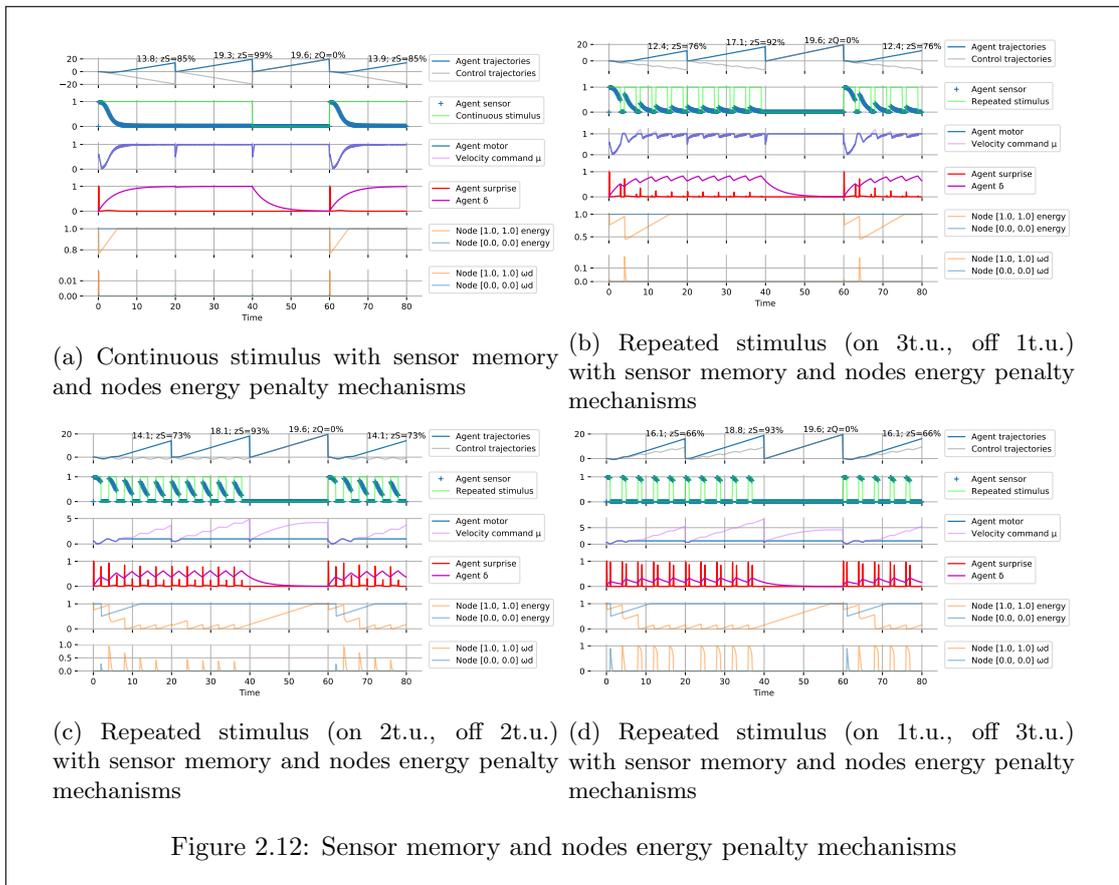
In this section, we first discuss some differences between our models. Then we give our thoughts about the characteristics of habituation not covered in the result section and whose study could be interested along two lines: either for modelers who would like to extend our work or because those properties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

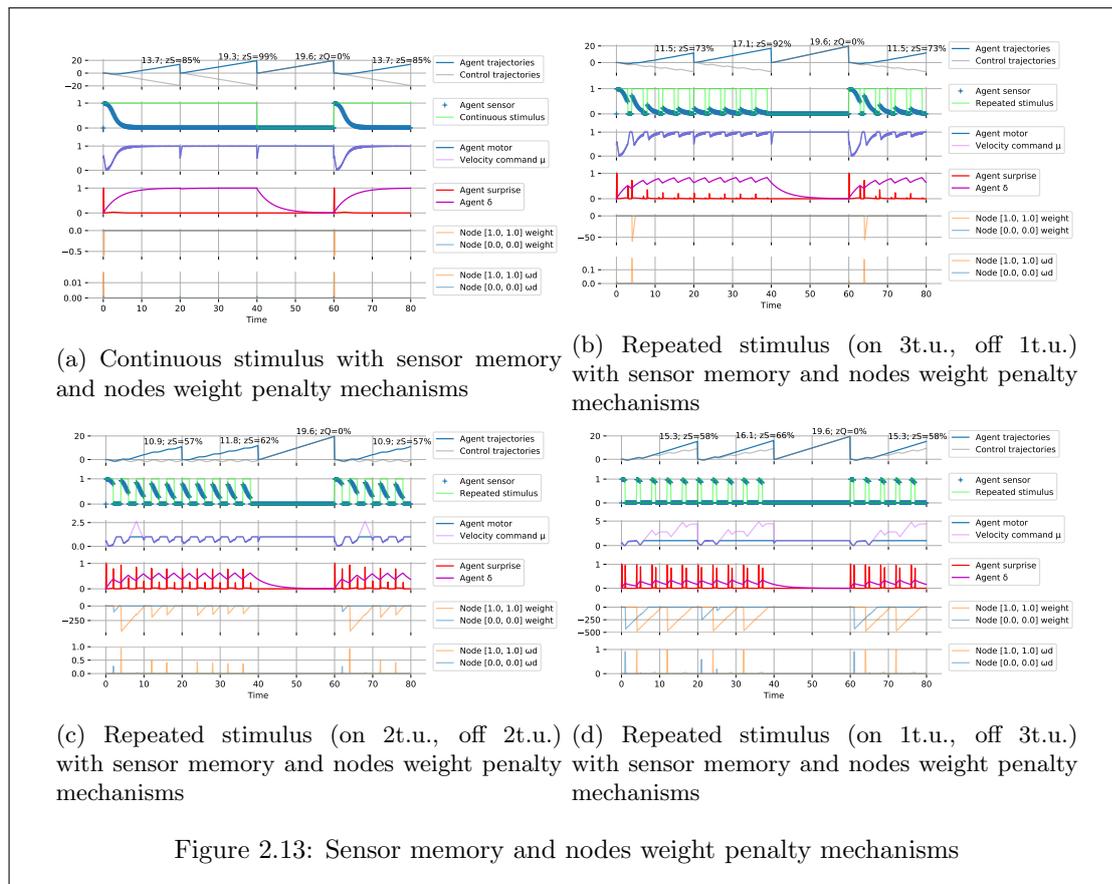
### 2.5.1 Models comparison

#### 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. 2.1), 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 imply 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 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 2.4.2, 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.

### Nodes energy vs nodes weight

There are also notable differences between the two nodes penalty models.

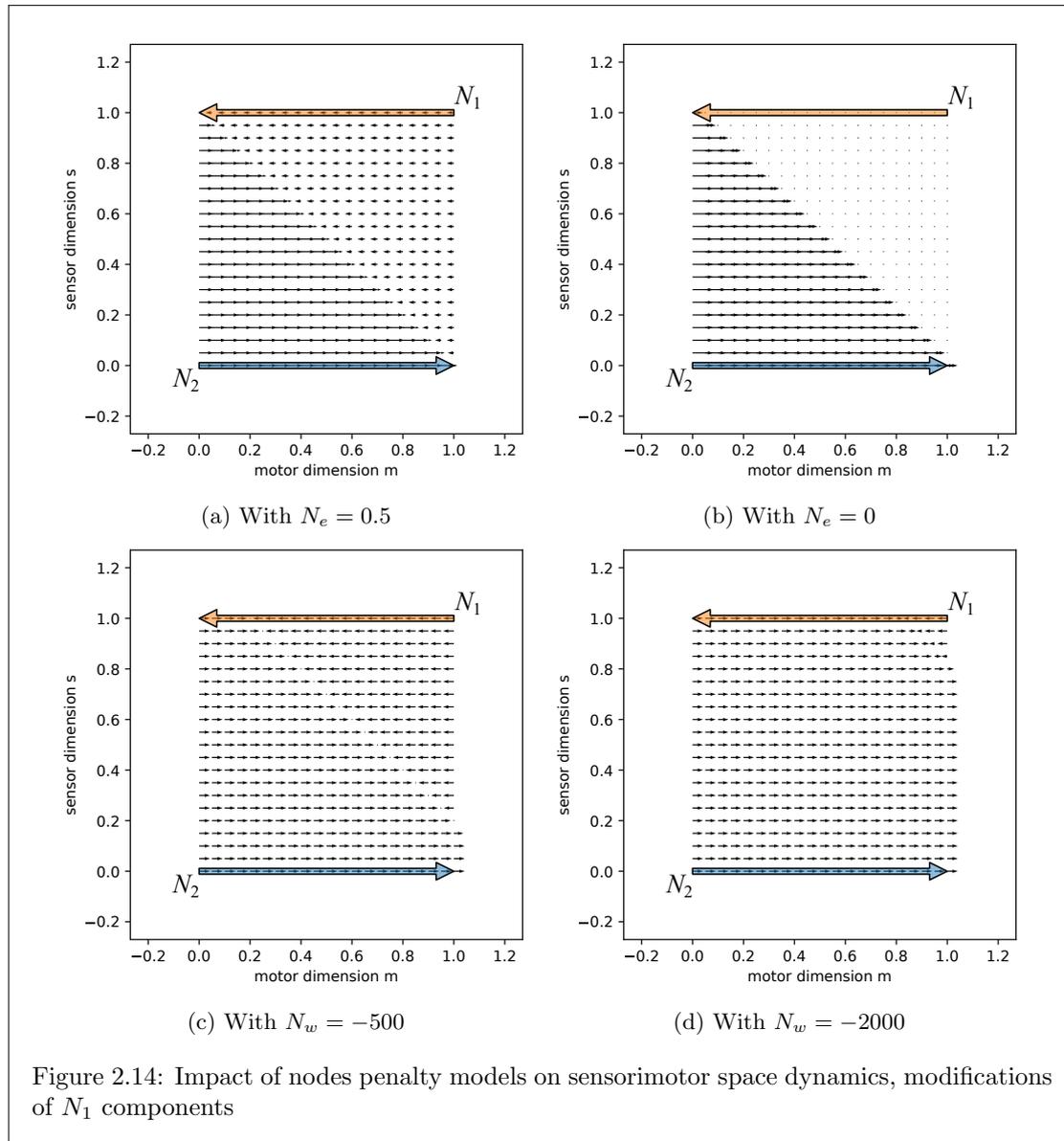
The main difference is that in the model impacting nodes weight, a lower weight means a lower  $\omega$  (see equation (1.6)). One of the consequences is that a decrease in  $\omega$  leads to a decrease of the impact of the surprise at a given distance (see eqs. (2.11) and (2.12)). Therefore, the lower the node weight, the harder it is to inhibit it further. 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  $\mathbf{x}_{SM}$  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. 2.5a and 2.6a).

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. 2.14a and 2.14b). With decreasing weight, the node command stays the same, but the border between nodes influences shifts in favor of the strongest node (see figs. 2.14c and 2.14d).

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.



## 2.5.2 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.

### 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.

### 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.

### 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. (1.2)), the velocity command  $\mu$  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.

### 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.

### 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.

**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.

**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.



## Chapter 3

# Sensitization Models

### 3.1 Introduction

What drives the behavior of an organism? That fundamental question has been addressed in the literature for a variety of organisms and through different angles. For some species, part of the answer lies in the associative learning mechanisms but there also exists more widespread non-associative learning mechanisms (Alex and Rankin, 2017; Byrne and Hawkins, 2015; Eisenstein et al., 1982). Habituation and sensitization, the first translating behaviorally in a response decrease while the second in a response increase, are two such mechanisms that are so widely observed in the phylogeny that it suggests their fundamental role in behavioral learning (Eisenstein et al., 2012; McSweeney and Murphy, 2009).

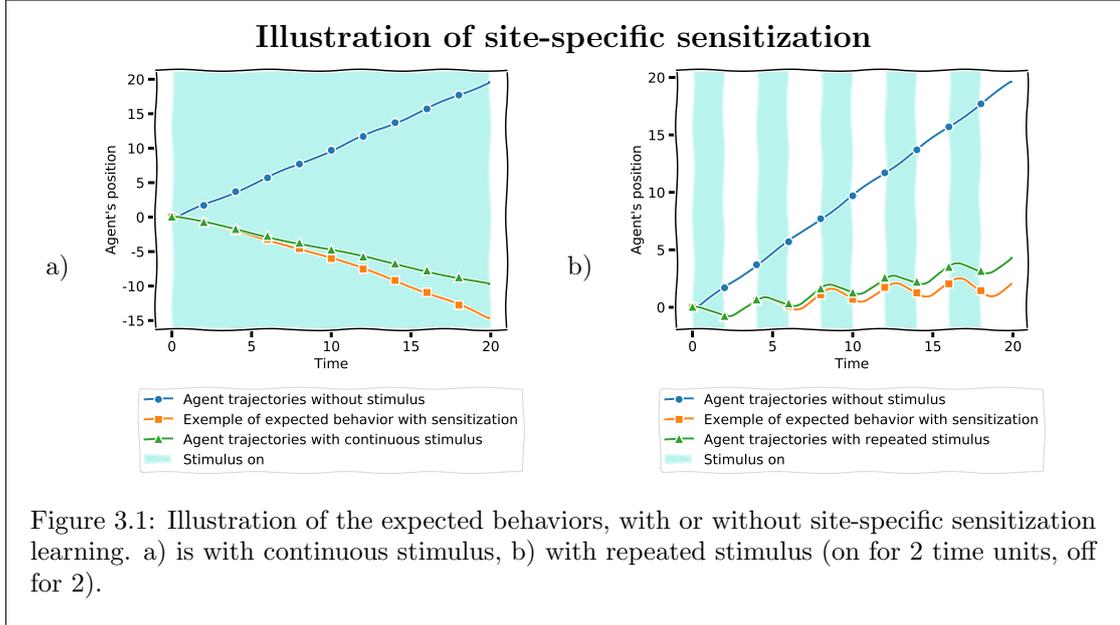
Sensitization is defined as either the enhancement of the response to a repeated stimulus or as a generalization of this response to another stimulus. Hence, sensitization, rather than habituation, appears as an interesting mechanism to investigate how an agent can broaden its behavioral responses. We describe in more details sensitization in section 3.2.

The objective of this work is thus to propose computational models *inspired* by sensitization and analyze their properties. In a future work, we will explore how these models can help artificial agents acquire and develop new behaviors.

To our knowledge, few works have addressed computational models of non-associative learning. In (Hong et al., 2020), both habituation and sensitization are modeled on neurons for a neuromorphic processor. A weak, repeated stimulus leads to habituation of the neuron activity, whereas a strong stimulus leads to the sensitization of the neuron activity. They show that such architecture lead to more robust hardware with better aging properties.

In this paper, we propose new computational mechanisms inspired by sensitization 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. 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 three extensions to the IDSM framework inspired by sensitization. One is based on the idea of having a *sensor memory* which amplifies the perception of the stimulus. The second approach we explore is based on the idea that *surprising events* trigger



sensitization by increasing the amplitude of response that would be normally triggered. The last model is centered on the response generalization to other stimuli. This paper is structured as follows. We start by recalling the biological definition of sensitization and its behavioral correlates in section 3.2. The Iterant Deformable Sensorimotor Medium (IDSMM) framework is then described in section 1.1. The proposed models are detailed in section 3.3 and various experiments are presented and analyzed in section 3.4. We conclude the paper with a discussion on our models.

## 3.2 Sensitization

Sensitization is a *non-associative* learning observed within numerous and various species. Sensitization is defined as the enhancement of a response behavior after the presentation of some stimulus and can be separated in two forms: *site-specific* sensitization and *pseudo-conditioning* (Byrne and Hawkins, 2015).

By definition, site-specific sensitization is the enhancement of the behavioral response following repeated presentation of the same stimulus (see fig. 3.1). The other form of sensitization, pseudo-conditioning, is a form of learning allowing a behavioral response to a given *unconditioned* stimulus  $US$  to generalize, temporarily, as a response to another *conditioned* stimulus  $CS$  (Hawkins and Kandel, 1984). The difference with the associative learning *classical conditioning* is that the *conditioned* response occurs without having to present the  $US$  and the  $CS$  at the same time.

### 3.3 Models

#### 3.3.1 Site-specific sensitization

In order to integrate site-specific sensitization as defined in section 3.2 into the IDSM, we propose two models. In the first one, a sensor memory enhances the perception of the stimulus. In the other, site-specific sensitization occurs through the reinforcement of nodes influence through weight modification.

For the experiments on these two models, we build an IDSM with two nodes as shown on Figure 3.2. Our objective with these two nodes is to produce a default behavior (without sensitization) of moving upward without stimulus and downward when the stimulus is on. These nodes are defined by:

- $N_1 = \langle \mathbf{p} = [1, 1], \mathbf{v} = [-1, 0], w = -2000 \rangle$ ,
- $N_2 = \langle \mathbf{p} = [0, 0], \mathbf{v} = [1, 0], w = -2000 \rangle$ .

If the agent perceives the stimulus (the sensory response is  $s = 0.75$ ), 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. Furthermore, as seen from figure 3.2, the more intense or strongly perceived the stimulus (*i.e.* high sensor value), the more the motor command tends to diminish the motor value. For example, with a sensor value of 0.75, the normalized motor will stabilize on 0.25 and the agent will go down slowly, and with a sensor value of 0.5, the normalized motor will stabilize on 0.5 and the agent will stop moving.

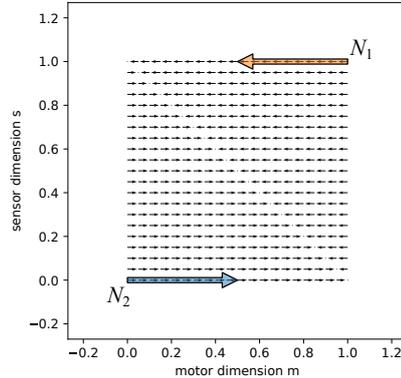


Figure 3.2: IDSM with two nodes used in site-specific sensitization experiments. See text for details.

#### Site-specific sensitization through sensor memory

Site-specific sensitization is defined as an increase in response to a repeated stimulus. In the first model we propose, a sensor memory  $\delta$  integrates the value of the sensor and increases virtually the agent's perception of the stimulus. This increase in perception ultimately causes an increase in the expression of the response. The sensor memory  $\delta$  keeps track of the evolution of the perceived stimulus. In this model, the sensor value seen by the IDSM is the *sensitized* sensor value, which is modulated by a function (sensitization factor *sen*) of the difference between the sensor value and its mean over a recent time window  $\delta$ . The higher the sensor memory  $\delta$  is, the more the motor response is pronounced, without altering the nodes in the IDSM. If the

stimulus is switched off, the sensor memory  $\delta$  will decrease and then the motor response recovers progressively.

The sensor memory  $\delta$  is computed as a low-pass filter over the *non-sensitized* sensor value  $nss$ :

$$\frac{\partial \delta}{\partial t}(\mathbf{x}_E, t) = \beta(nss(\mathbf{x}_E) - \delta(\mathbf{x}_E, t)) \quad (3.1)$$

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 sensitized sensor value  $s_{sen}$ , used as the sensor component of the agent sensorimotor position  $\mathbf{x}_{SM}$ , is computed from the sensor value  $nss$  and the sensor memory  $\delta$  as:

$$s_{sen}(\mathbf{x}_E, t) = nss(\mathbf{x}_E)(1 + \gamma(1 - sen(\mathbf{x}_E, t))) \quad (3.2)$$

$$sen(\mathbf{x}_E, t) = \frac{1}{1 + e^{\alpha(\|\delta(\mathbf{x}_E, t) - nss(\mathbf{x}_E)\| - 0.5)}} \quad (3.3)$$

with  $\gamma = 0.25$  and  $\alpha = -10$ . From the plot of the sensitization factor  $sen$  on Figure 3.3, the stimulus is perceived with little or no change if its value is far from the memory (large  $\|\delta - nss\|$ ) and enhanced according to  $\gamma$  if its value is close to the sensor memory ( $\delta \approx nss$ ). Therefore, if a stimulus is present for some times, the perceptual response for the agent will be increased and its behavioral response will be increased as well which produces the expected sensitization.

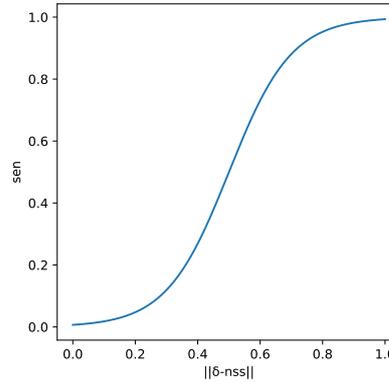


Figure 3.3: The sensitization factor  $sen$  as a function of the difference between the sensor value  $nss(\mathbf{x}_E)$  and the sensor memory  $\delta$ .

### Site-specific sensitization through impact on nodes weight

While the previous model globally affects the sensory perception, it is also possible to implement site-specific sensitization by local mechanisms, at the node level in the IDSM by modifying their weights.

The following model is based on the concept of surprise in the sensorimotor space: each time the agent is “surprised” by its actual position in the sensorimotor space, the influences of the surrounding nodes are reinforced by increasing their weights. The more often a node is reinforced, the more influence it gets on the IDSM dynamics. To allow the gradual recovery of the non-sensitized response, the influence of nodes continuously decreases over time.

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

$$sur(t) = \|\mathbf{x}_{SM}(t) - \hat{\mathbf{x}}_{SM}(t)\| \quad (3.4)$$

$$\hat{\mathbf{x}}_{SM}(t) = \mathbf{x}_{SM}(t - \Delta t) + [\boldsymbol{\mu}, 0]\Delta t \quad (3.5)$$

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.

This surprise is then used to modulate the weights of the nodes. The weight value increases when there is surprise  $sur$  and recovers (*i.e.* decreases) gradually over time. This model is more in line with the original IDSM.

$$g_w(N, t) = \beta_w \cdot sur(t) \cdot d(N_{\mathbf{p}}, \mathbf{x}_{SM}) \quad (3.6)$$

$$N_w(t + \Delta t) = N_w(t) + g_w(N, t) - \gamma_w \Delta t \quad (3.7)$$

with weight gain value  $\beta_w = 1000$  and weight recovery value  $\gamma_w = 75$ . The nodes weights start at  $-2000$  and are kept in  $[-2000, \infty[$ .

Hence, for a given node  $N$ , its weight  $N_w$  is reinforced when the agent sensorimotor position  $\mathbf{x}_{SM}$  is close to the node position  $N_{\mathbf{p}}$  and the surprise is not 0. Such a surprise occurs only for event linked to activation or deactivation of the stimulus.

### 3.3.2 Pseudo-conditioning through generalization of nodes

Pseudo-conditioning is a form of learning allowing a behavioral response to a given stimulus to be expressed as response to another stimulus, a sort of generalization, to other stimuli, of this response. The following mechanism aims at extending the response dynamics from one sensory dimension of the IDSM to another.

For the experiments on this model, there are three dimensions in the sensorimotor space  $SM$ : one motor dimension and two sensor dimensions. The nodes in the sensorimotor space  $SM$  are shown on Figure 3.4, without representation of the second sensor dimension for simplicity (the second sensor value is  $s_2 = 0$ ). The two nodes are defined by:

- $N_1 = \langle \mathbf{p} = [0.5, 1, 0], \mathbf{v} = [-0.5, 0, 0], w = 0 \rangle$ ,
- $N_2 = \langle \mathbf{p} = [0, 0, 0], \mathbf{v} = [0.5, 0, 0], w = 0 \rangle$ .

We add a mask component  $\mathbf{N}_m$  of dimension  $K$  with components in  $[0, 1]$  to every node  $N$ .  $\mathbf{N}_m$  is used to compute a virtual *masked* distance  $dm$ , between the node  $N$  and the agent sensorimotor position  $\mathbf{x}_{SM}$ :

$$dm(N_{\mathbf{p}}, \mathbf{x}_{SM}) = \frac{2}{1 + \exp(k_d \|N_{\mathbf{p}} - \mathbf{x}_{SM}\|_m^2)} \quad (3.8)$$

$$\|N_{\mathbf{p}} - \mathbf{x}_{SM}\|_m^2 = \sum_i \mathbf{N}_{m_i}^2 (N_{\mathbf{p}_i} - \mathbf{x}_{SM_i})^2 \quad (3.9)$$

This distance  $dm$ , which artificially decreased the sensorimotor distance between the agent and the nodes, is used instead of  $d$  in equations 1.3.

It follows that for a given dimension  $i$ , the most specific the associated mask component can be is  $\mathbf{N}_{m_i} = 1$  (*i.e.* the distance between  $N_{\mathbf{p}_i}$  and  $\mathbf{x}_{SM_i}$  is fully taken into account). And the

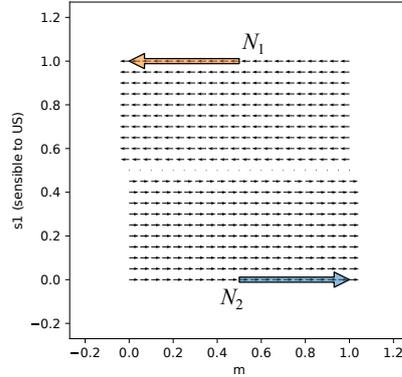


Figure 3.4: IDSM with two nodes used in pseudo-conditioning experiments. See text for details.

most general the associated mask component can be is  $\mathbf{N}_{m_i} = 0$  (*i.e.* the distance between  $N_{p_i}$  and  $\mathbf{x}_{SM_i}$  is completely ignored, as if  $N_{p_i} = \mathbf{x}_{SM_i}$ ).

A node starts with its mask equal to  $\mathbf{1}$ , the most specific possible. Then the mask is modified following:

$$f_{gen} = (\sqrt{K^\sigma} - \|\mathbf{x}_{SM}\|^\sigma - \|N_p\|^\sigma) \beta_m \omega d_{mN} \quad (3.10)$$

$$\frac{\partial \mathbf{N}_m}{\partial t}(\mathbf{x}_{SM}, t) = \gamma_m \cdot \mathbf{1} - f_{gen} \cdot \mathbf{N}_m \quad (3.11)$$

with  $\beta_m = 3$  and  $\gamma_m = 0.05$ . Each of the mask components is kept between 0 and 1.

Thus, the smaller the masked distance between a node  $N$  and the sensorimotor position of the agent  $\mathbf{x}_{SM}$ , the more the mask of the node  $\mathbf{N}_m$  is decreased (and so becomes generalized). Specialization (*i.e.* increase) of the mask occurs slightly over time. Parameters  $\beta_m$  and  $\gamma_m$  adjust respectively the generalization and the specialization of the masks.

## 3.4 Experiments and results

### 3.4.1 Site-specific sensitization setup

In order to measure if our models are capable of site-specific sensitization (*i.e.* produce an increased response to the continuous or repeated presentation of a single stimulus), we use the following protocol.

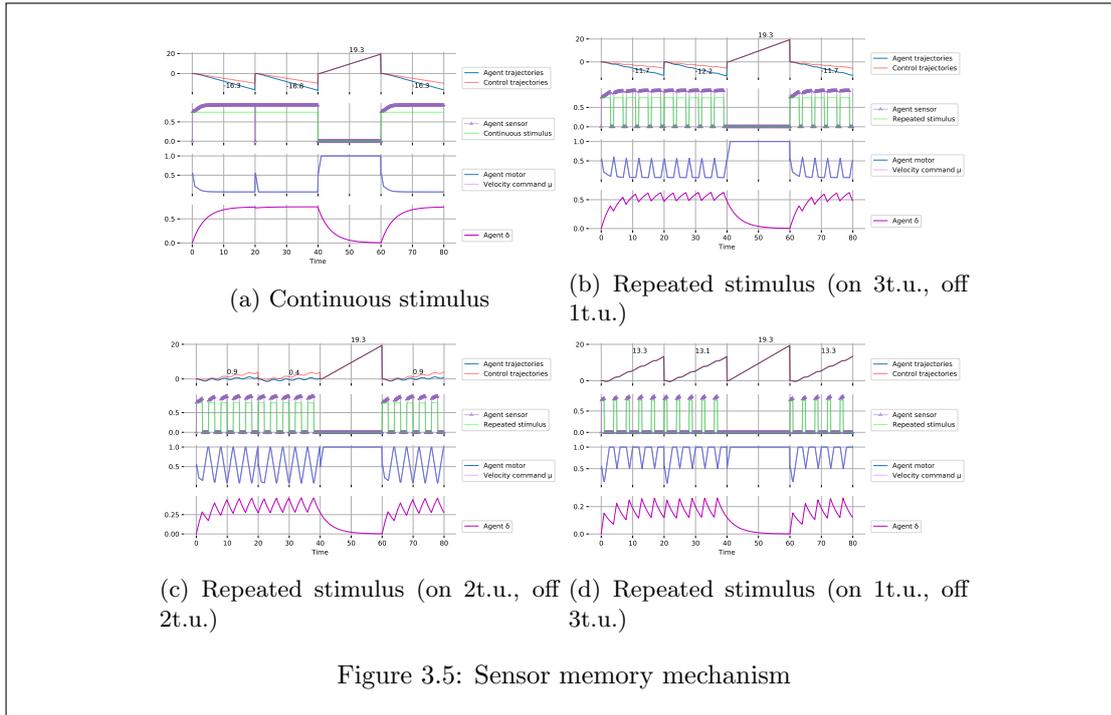
The experiment is split into different phases: stimulation phases (with continuous or repeated stimulus) and quiet phases (without any stimulus). First there are two stimulation phases to evaluate sensitization, and then a quiet phase followed by a new stimulation phase to evaluate the recovery of the response in time. 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  $\mathbf{x}_{SM} = [0.5, 0]$ . 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, and the stimulus value is 0.75 (necessary for the sensor model to virtually increase the sensor value).

- (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.

### 3.4.2 Site-specific sensitization results

Site-specific sensitization is 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 3.5-3.6.



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, the number indicated is the agent last position  $x_E$  for this phase.

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  $x_{SM}$ ). The third plot from the top shows the normalized motor value of the agent in the sensorimotor space (the first component of  $x_{SM}$ ), and the velocity command  $\mu$ . The last plots give more information depending on the model (sensor memory, energy of the nodes, weights of the nodes).

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).

For the sensor memory model (section 3.3.1), as shown on figure 3.5, longer stimulus activation produces stronger sensitization. This can be explained by the fact that the sensor memory  $\delta$

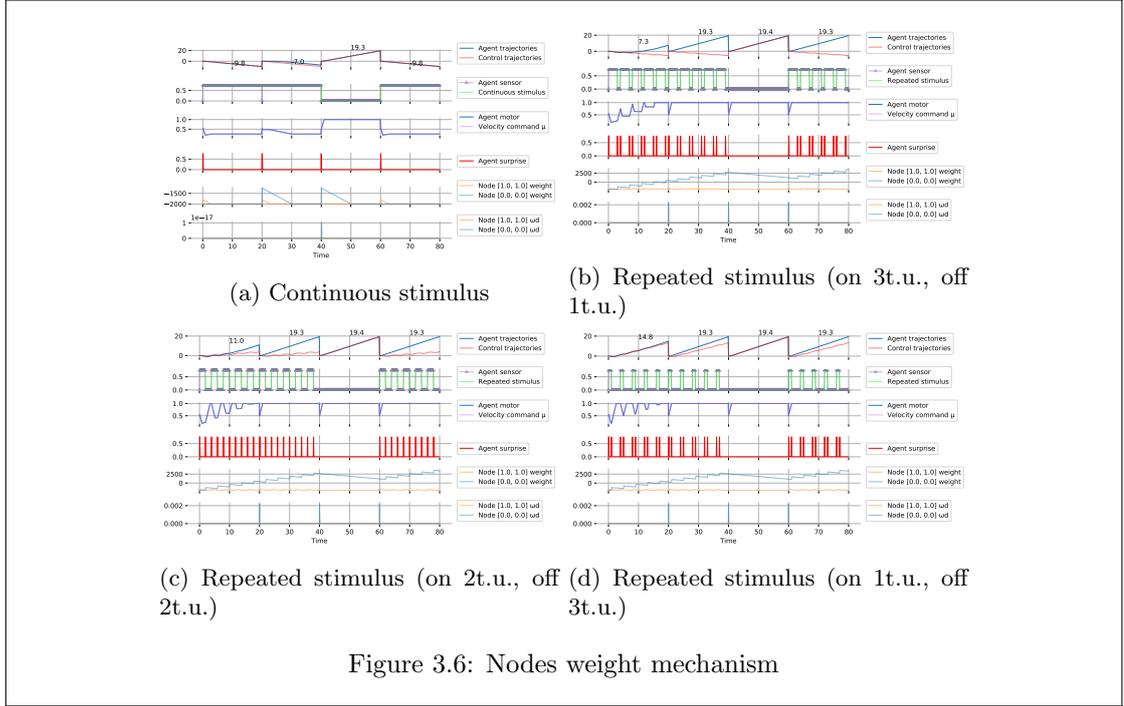


Figure 3.6: Nodes weight mechanism

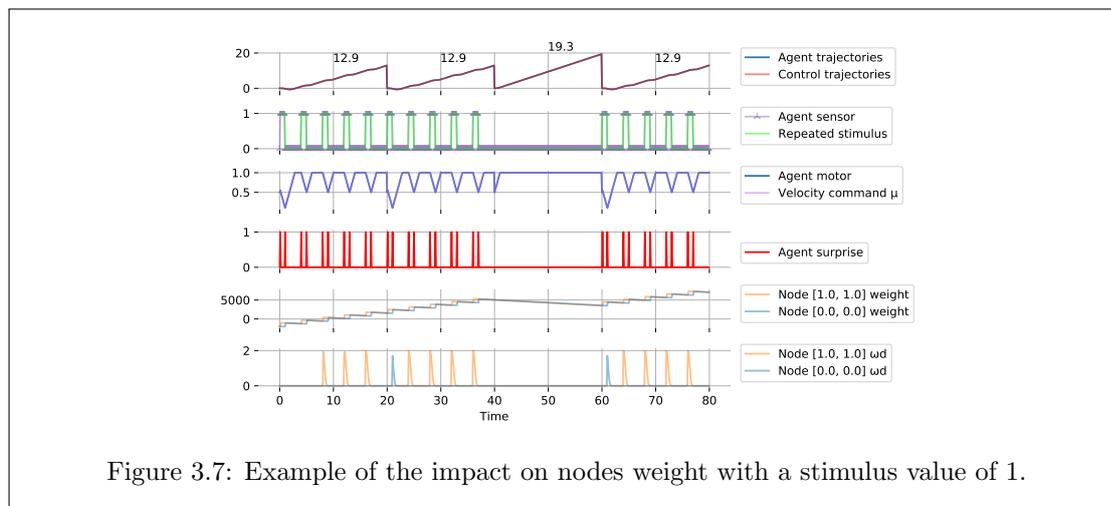
approximates the mean of the stimulus activation (*i.e.* 0.75, 0.5625, 0.375 and 0.1875 respectively with stimulation patterns (a), (b), (c) and (d), see bottom plots of the four subplots), hence its enhancing effect is accordingly stronger with longer activation time (see equation (3.3) and figure 3.3).

However, for the nodes weight model (see figure 3.6), the response (downward motion) is only seen with continuous stimulus (pattern (a)). There is a kind of “reverse” sensitization that reinforces the upward motion as it happen when the stimulus is absent. This is explained by the choice of the stimulus value of 0.75, too far from node  $N_1$  to have an significant impact on its weight. In the case where the value of the stimulus would be 1, the impact on the two nodes are similar, balancing each other influence and leading to no visible changes in the agent behavior compared to the control agent (see figure 3.7).

### 3.4.3 Pseudo-conditioning setup

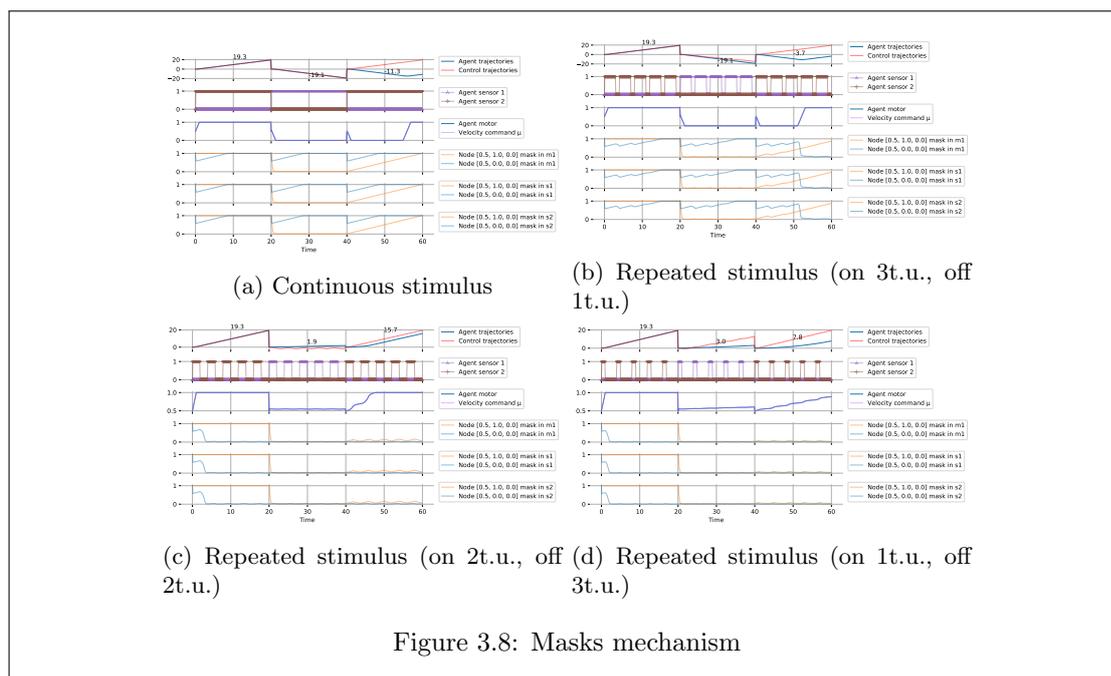
For the experiments on pseudo-conditioning, we need two stimuli: the unconditioned stimulus  $US$ , and the conditioned stimulus  $CS$ . The experiment is split into three phases of twenty time units. During the first phase, the  $CS$  is presented, to observe the agent behavior prior to pseudo-conditioning. The  $CS$  is then replaced by the  $US$  during the second phase. Lastly, the  $CS$  is presented again during the third phase.

As explained in section 3.3.2, the IDSM is modified accordingly. When either the  $CS$  or  $US$  is active, its value is set to 1. All other conditions are the same as in the experiments of section 3.4.1.



### 3.4.4 Pseudo-conditioning results

The results for pseudo-conditioning can be seen on figure 3.8.



As for site-specific sensitization experiments, the top plot displays the learning agent trajectory and the control agent trajectory, with the indication of the agent last position  $x_E$  for each phase.

The second plot from the top gives information regarding perception of both the stimuli by the agent. The third plot from the top shows the normalized motor value of the agent in the sensorimotor space (the first component of  $x_{SM}$ ), and the velocity command  $\mu$ . The three last

plots show the evolution of the nodes masks, one plot by dimension (motor, first sensor sensible to the  $US$ , and second sensor sensible to the  $CS$ ).

For the control agent, the default behavior (upward motion) is observed during the first and third phase, with presentation of the  $CS$ . During the second phase, with presentation of the  $US$ , a response behavior (downward motion) is observed.

With all patterns, the behavior of the learning agent during the third phase (*i.e.* presentation of the  $CS$  following presentation of the  $US$ ) is a mix between its behaviors of the first and second phases, starting as a “ $US$  influenced” behavior and ending as a “ $CS$  influenced” behavior (default behavior). In 3.8a, pseudo-conditioning has no influence in first and second phases. The pseudo-conditioning occurs as expected in the last phase. The evolution of the masks shows that only the node  $N_1$ , responsible for the response to the  $US$ , is generalized. The node  $N_2$ , responsible for default behavior, only generalizes slightly and temporarily following the reset of the agent at the start of each phase. Conversely, in all others patterns (see figs. 3.8b to 3.8d), the node  $N_2$  generalizes each time there is no active stimulus. In the pattern **(b)**, with the shorter deactivation time between stimulus activation, the learning agent in the second phase does not respond to deactivation of the stimulus since  $N_1$  stays more generalized than  $N_2$  and behaves as if the  $US$  was activated continuously during this phase. However, in third phase, pseudo-conditioning is shorter than in pattern **(a)**. In the patterns **(c)** and **(d)**, with the longer deactivation time, generalization of node  $N_2$  leads to a tie between the two nodes influences, which results in weaker motor commands and extends the time needed for the agent to move.

### 3.5 Discussion about Sensitization

We have proposed two models (sensor memory model and impact on nodes weight model) to provide site-specific sensitization to an agent based on the IDSM, and one model for pseudo-conditioning (with nodes mask).

The sensor memory model does show some sensitization, the amplification of sensor value allowing the agent to be more influenced by the node responsible of the response behavior.

The nodes weight model also show some sort of sensitization in the sense that it is the default behavior (upward motion) that is more expressed. This could be defined as “reverse sensitization”.

For pseudo-conditioning, the generalization of the nodes through the masks results in a transfer of the behavioral response from the stimulus initially triggering it to other stimuli.

Note that the results depend to some extent on the positions of the nodes and the values of parameters. Thus it is quite difficult to have a precise idea of what the models could give in a more complex IDSM, in particular with lots of nodes and with nodes generation in the low nodes density zones of the sensorimotor space  $SM$ , as done in the original IDSM.

The results of both site-specific sensitization and pseudo-conditioning are encouraging for future works. This opens the road toward proposing mechanisms for an agent to generalize behaviors to other stimuli. That would allow the agent to experiment known behaviors in new situations.

The mask mechanism we propose for pseudo-conditioning leads to a widespread generalization : once a node is generalized it impacts the behavior of the agent equally everywhere in the sensorimotor space. This generalized behavior is only counterbalanced by the existence of other nodes (through the usual IDSM equations). Maybe this “competition” between nodes is not sufficiently strong to preserve other existing behaviors. In such a case, it may be needed to introduce a specialization mechanism shaping the sensorimotor regions where the node should actually generalize. For example, specialization could be triggered by the discrepancy between

---

the component  $N_v$  of node  $N$  and the last movement of the agent in the sensorimotor space  $SM$ , since such difference would mean that the motor command stored in  $N$  is not in line with the current environment. Adding another non-associative learning like habituation (*i.e.* inhibition of the response to a repeatedly presented stimulus) could also help prevent over-generalization. The dual-process theory (Groves and Thompson, 1970) and the behavioral homeostasis theory (Eisenstein et al., 2001, 2012) are works studying how non-associative learning fit together and what could be the impacts of such association.



# Bibliography

- Alex, J. Y. and Rankin, C. H. (2017). Nonassociative learning in invertebrates. In *The Oxford Handbook of Invertebrate Neurobiology*.
- 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)*, pages 679–684. IEEE.
- Byrne, J. H. and Hawkins, R. D. (2015). Nonassociative learning in invertebrates. *Cold Spring Harbor perspectives in biology*, 7(5):a021675.
- Egbert, M. D. and Barandiaran, X. E. (2014). Modeling habits as self-sustaining patterns of sensorimotor behavior. *Frontiers in human neuroscience*, 8:590.
- Eisenstein, E., Brunder, D., and Blair, H. (1982). Habituation and sensitization in an aneural cell: Some comparative and theoretical considerations. *Neuroscience & Biobehavioral Reviews*, 6(2):183–194.
- Eisenstein, E., Eisenstein, D., and Smith, J. C. (2001). The evolutionary significance of habituation and sensitization across phylogeny: A behavioral homeostasis model. *Integrative Physiological & Behavioral Science*, 36(4):251–265.
- Eisenstein, E. M., Eisenstein, D. L., Sarma, J. S., Knapp, H., and Smith, J. C. (2012). Some new speculative ideas about the “behavioral homeostasis theory” as to how the simple learned behaviors of habituation and sensitization improve organism survival throughout phylogeny. *Communicative & Integrative Biology*, 5(3):233–239.
- Groves, P. M. and Thompson, R. F. (1970). Habituation: a dual-process theory. *Psychological review*, 77(5):419.
- 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.
- Hawkins, R. D. and Kandel, E. R. (1984). Is there a cell-biological alphabet for simple forms of learning? *Psychological review*, 91(3):375.
- 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, T. A., Yu, A. J., and Rankin, C. H. (2019). Habituation is more than learning to ignore: multiple mechanisms serve to facilitate shifts in behavioral strategy. *BioEssays*, 41(9):1900077.
- McSweeney, F. K. and Murphy, E. S. (2009). Sensitization and habituation regulate reinforcer effectiveness. *Neurobiology of learning and memory*, 92(2):189–198.
- Randlett, O., Haesemeyer, M., Forkin, G., Shoenhard, H., Schier, A. F., Engert, F., and Granato, M. (2019). Distributed plasticity drives visual habituation learning in larval zebrafish. *Current Biology*, 29(8):1337–1345.
- Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., Coppola, G., Geyer, M. A., Glanzman, D. L., 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, C. H., Beck, C. D., and Chiba, C. M. (1990). *Caenorhabditis elegans*: a new model system for the study of learning and memory. *Behavioural brain research*, 37(1):89–92.
- Stanley, J. C. (1976). Computer simulation of a model of habituation. *Nature*, 261(5556):146–148.
- Thompson, R. F. and Spencer, W. A. (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, E. H., Ponce, F. P., and Wagner, A. R. (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.

*Inria*

**RESEARCH CENTRE  
NANCY – GRAND EST**

615 rue du Jardin Botanique  
CS20101  
54603 Villers-lès-Nancy Cedex

Publisher  
Inria  
Domaine de Voluceau - Rocquencourt  
BP 105 - 78153 Le Chesnay Cedex  
[inria.fr](http://inria.fr)

ISSN 0249-6399